8-2001

Positional cloning of the psrt mutations on mouse chromosome 7

Lisa Smith Webb
University of Tennessee

To the Graduate Council:

I am submitting herewith a dissertation written by Lisa Smith Webb entitled "Positional cloning of the psrt mutations on mouse chromosome 7." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biomedical Sciences.

Dabney K. Johnson, Major Professor

We have read this dissertation and recommend its acceptance:

Edward J. Michaud III, Rebecca A. Prosser, Gary A. Sega

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Lisa Smith Webb entitled "Positional Cloning of the *psrt* Mutations on Mouse Chromosome 7." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biomedical Sciences.

Dabney K. Johnson, Major Professor

We have read this dissertation
and recommend its acceptance:

Edward J. Michaud, III

Rebecca A. Prosser

Gary A. Sega

Accepted for the Council:

Interim Dean of The Graduate School

POSITIONAL CLONING OF THE *psrt* MUTATIONS ON MOUSE CHROMOSOME 7

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Lisa Smith Webb
August, 2001

DEDICATION

This dissertation is dedicated

to my parents, Homer Leslie Smith and Barbara Jean Henry Smith;

to my husband, Rick William Webb;

to my son, Christopher Matthew Webb;

and in memory of my grandmother,

Ruby Ruth McNeilly Smith,

whom I have never stopped loving or missing.

Thank you for your unconditional love, your unfailing support,

and your encouragement.

This dissertation is yours as much as it is mine.

Laudate Dominum.

ACKNOWLEDGEMENTS

I would like to thank Dr. Dabney K. Johnson, my Committee Chairperson and Major Professor, for giving me the support and encouragement to undertake and complete such a fascinating and challenging project. I would also like to thank Dr. Gary A. Sega who, in addition to sitting on my Committee, directed the bioanalytical work in my project. I am very grateful to my other Dissertation Committee Members, Dr. Edward J. Michaud, III, and Dr. Rebecca A. Prosser, as well as to Dr. Eugene M. Rinchik and Dr. Liane B. Russell, who were gracious enough to sit in on my Committee meetings and read my dissertation. They were always available to listen and willing to offer guidance and suggestions. My project and the resulting dissertation would not have been complete without their contributions.

I would like to thank Dr. Madhu S. Dhar for her friendship and guidance, and for taking the time to teach me molecular biology. I am also indebted to Dr. Cymbeline T. Culiat, who arranged to have the mutation scanning work performed on the *Prmt3* cDNA. I owe a tremendous debt of gratitude to the many dedicated scientists, technicians, and support staff at the ORNL Mouse House. They were always willing to help me with whatever I needed, no matter how busy they were, and I feel blessed to count them as my friends. And last of all, I would like to thank the UT-ORNL Graduate School of Genome Science and Technology for its financial support. I am especially grateful to Dr. Jeffrey M. Becker for his encouragement and sage advice and to Ms. R. Kay Gardner for her friendship, support, and assistance.

ABSTRACT

The *psrt* mutations, 723SJ and 1060SJ, are ENU-induced, non-complementing mutations that produce identical phenotypes and map to the same *p*-deletion interval on mouse chromosome 7. *psrt* stand for profound seizure and runting; animals exhibiting this phenotype are runted and have severe seizures that are first detected at seven to ten days of age. Homozygous and hemizygous mutants typically live 15 to 18 days. A positional cloning strategy was employed to identify the gene responsible for this phenotype. The region of the genome containing *psrt* was better defined by mapping molecular markers and determining the breakpoints of $p^{3RD300H}$, a *p*-deletion that does not complement the phenotype. A physical map was constructed in the minimal deletion interval determined to contain the mutation, and a Bacterial Artificial Chromosome from the physical map was sequenced to identify candidate genes. Three genes were identified, two from sequence analysis (Tat-interacting protein (30 kDa), *Tip30*, and protein arginine *N*-methyltransferase 3, *Prmt3*) and one from a search of the genome sequencing databases from the human region of homology (glycine transporter type 2, *GLYT2*). No mutations were detected in the *Tip30* cDNA and *Glyt2* was not molecularly characterized. However, the 5' end of the 723SJ *Prmt3* transcript could not be amplified, indicating a possible chromosomal rearrangement or deletion, and mutation screening by temperature-gradient capillary electrophoresis was utilized to detect mismatches between the control and both the 723SJ and 1060SJ cDNAs. This indicates that mutations in the *Prmt3* transcript are present in both strains of mutants.

Initial sequence analysis failed to detect a mutation; however, further sequencing will be performed to verify the mismatches. Since *Prmt3* could be involved in the methylation of the single methylarginine residue in the myelin basic protein, an assay to detect the presence of this residue in the myelin basic protein hydrolysate was performed on mutants and controls from the 1060SJ strain of mice. Results indicate that the methylarginine is easily detected in the 1060SJ control animals, but could not be detected in the 1060SJ *psrt* mutants. The molecular and biochemical evidence suggests that *Prmt3* is the gene responsible for the *psrt* phenotype.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AdoMet	*S*-adenosylmethionine
A/T	adenine/thymine
BAC	bacterial artificial chromosome
bp	basepair
cDNA	complementary DNA
cM	centimorgan
CNS	central nervous system
DNA	deoxyribonucleic acid
ENU	*N*-ethyl-*N*-nitrosourea
EST	expressed sequence tag
eV	electron volts
GC/MS	gas chromatography/mass spectrometry
kcal/mol	kilocalories per mole
kb	kilobasepair
kDa	kilodalton
L	liter
μ	micro
m	milli
M	molar
MBP	myelin basic protein

Me-R	methylarginine
Mmu	mouse chromosome (*Mus musculus*)
mRNA	messenger RNA
MTBSTFA	*N*-methyl-*N*-(*tert*-butyldimethylsilyl)trifluoroacetamide
n	nano
nt	nucleotide
ORNL	Oak Ridge National Laboratory
P	*pink-eyed dilution* locus
PCR	polymerase chain reaction
PFGE	pulsed field gel electrophoresis
PNS	peripheral nervous system
PRMT	protein arginine methyltransferase
psi	pounds per square inch
psrt	profound seizure and runting
R	arginine
RACE	rapid amplification of cDNA ends
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rRNA	ribosomal RNA
RT-PCR	reverse transcription-polymerase chain reaction
ru2	*ruby-eye 2* locus
SAM	*S*-adenosylmethionine

SLT	Specific Locus Test
SNP	single nucleotide polymorphism
STS	sequence tagged site
t-BDMS	*tert*-butyldimethylsilyl
TGCE	temperature-gradient capillary electrophoresis
UTR	untranslated region

CHAPTER I

OVERVIEW OF THE PROJECT AND DESCRIPTION OF THE *psrt* PHENOTYPE

OVERVIEW

This dissertation documents the positional cloning of the gene responsible for the profound seizure and runting (*psrt*) phenotype on mouse chromosome 7. Chapter one contains an overview of the project and a description of the phenotype. Chapter two contains the background information on the genetic reagents utilized in this project as well as the methodology utilized to produce the mutants. Chapter three contains a summary of the *trans*-complementation data that allowed the mutation to be fine-mapped. Chapters four, five, seven and eight contain experimental data and include five major sections: abstract, introduction, materials and methods, results, and discussion. Chapter four covers physical mapping in the *psrt* region, chapter five covers the sequence analysis of genomic clones in the region, chapter seven covers an expression analysis of candidate genes and a sequence analysis of their transcripts, and chapter eight covers the analysis of methylarginine in the myelin basic protein of mutant and control animals. Chapter six consists of a detailed literature review of the candidate genes and Chapter nine is an evaluation of the candidates. Chapter ten, the final chapter, is a brief assessment of future work that is needed to give closure to the project. Literature references and an appendix containing the tables, figures and legends follow the ten chapters.

THE *psrt* PHENOTYPE

This project undertook the positional cloning of the gene responsible for a juvenile lethal, neurological phenotype called *psrt*, for "profound seizure and runting." As the name implies, this phenotype manifests as severe seizures and runting. The seizures, characterized by stiffened and extended limbs, are first detected in the mutants at around seven to ten days after birth. Seizure activity increases in intensity and duration until the animal eventually dies, at around 15-18 days. Mutants can easily be distinguished from their non-mutant littermates by their runted appearance, even at birth. Their development appears to be delayed, since not only are they smaller in size, but also their fur is sparse and slow to come in. This runtedness is not due to starvation because milk is visible in the pups' stomachs, even after the onset of seizure activity and up until the time of death.

CHAPTER II

BACKGROUND

INTRODUCTION

This research was performed using mutant animals generated over the years in the mutagenesis programs at the Oak Ridge National Laboratory (ORNL). What follows is a brief overview of the development of the genetic and molecular reagents utilized in these mutagenesis studies as well as a description of the mutagenesis programs themselves.

THE SPECIFIC LOCUS TEST (SLT)

The morphological specific locus test (SLT) was developed in the late 1940s by William L. (Bill) Russell at the ORNL to investigate the induction of heritable mutations in the mammalian germline by ionizing radiation (Russell, 1951). To enable the recovery of new recessive mutations, Russell constructed the mouse T (tester) stock, which is homozygous for then-existing recessive alleles at seven loci that code for visible phenotypes in the mouse. The seven loci, located on five chromosomes, include six that affect coat color and one that affects ear morphology. The recessive T-stock alleles at three unlinked loci are: non-agouti (*a*) at the *a* locus on chromosome 2; brown (*b*; *Tyrp1*^{b}) at the tyrosine-related protein (*Tyrp1*) locus on chromosome 4; and piebald spotting (*s*; *Ednrb*^{s}) at the endothelin receptor type B (*Ednrb*) locus on chromosome 14. Alleles at two sets of linked loci are: pink-eyed dilution (*p*) at the *p* locus and chinchilla

(c^{ch}; $Tyr^{c\text{-}ch}$) at the Tyrosinase (*Tyr*) locus, 15 centimorgans (cM) apart on chromosome 7; and dilute (*d*; $Myo5a^{d}$) at the myosin 5A (*Myo5a*) locus and short-ear (*se*; $Bmp5^{se}$) at the bone morphogenetic protein 5 (*Bmp5*) locus, 0.16 cM apart on chromosome 9 (Russell, 1951).

The SLT is performed by exposing mice that are homozygous wild type at all seven specific loci to a potential mutagen, then mating them to T-stock animals. If no new mutations are induced at any of the seven loci, all offspring of this mating are phenotypically wild type; however, if one or more recessive mutations were generated at any of the seven loci, mutant animals would be detected in the first-generation of offspring. Dominant visible mutations induced anywhere in the genome are detectable. If a newly induced mutation at one of the specific loci is lethal, this new mutation can still be recovered in lines derived from the heterozygous primary F_1 mutant by appropriate breeding, and can be perpetuated in a stock. The beauty of the SLT is threefold: since the loci affect coat color and ear morphology, the mutants are easily, quickly, and objectively recognized and categorized by their physical appearance; the mutations can be detected in only one generation; and all mutations, including any that are recessive lethals, can be recovered and perpetuated (Russell, 1951).

The SLT was initially used at the Oak Ridge National Laboratory to ascertain the rate of induction of heritable mutations by ionizing radiations in mouse germ cells, and, specifically, the effects of various physical and biological variables on the frequency and nature of the induced mutations. This information was desperately needed, since the effects of ionizing radiation on the human germ line were unknown at

the time. The only other organisms in which radiation-mutagenesis studies had been performed at that time were maize, *Zea mais*, and the fruit fly, *Drosophila melanogaster*; however, Russell felt that data from mammalian species were necessary because they would more closely estimate the effect on humans. He designed the SLT primarily to give quantitative data by measuring the mutation rate at each of the seven loci. However, he also designed the SLT to give qualitative information about the mutations emerging from the studies, since two of the loci are very tightly linked (*d* and *se* are a mere 0.16 cM apart on chromosome 9) and a point mutation or intragenic lesion would affect only one of the two loci while a deletion might affect both loci (Russell, 1951).

Because the SLT measured mutation rates at each of the specific loci, Russell was able to conclude early in his work that genes vary greatly in their mutability (Russell, 1951). Bill and Liane B. (Lee) Russell determined that chronic doses of radiation produced fewer detectable mutations in spermatogonial stem cells (Russell *et al.*, 1958) and in oocytes (reviewed in Russell and Russell, 1992) than the same amount of radiation administered in a single, acute dose, providing early evidence for repair of premutational damage (Russell *et al.*, 1958). They also used the SLT to study the genetic effects of numerous chemical mutagens. These experiments revealed that the stage at which germ-cells were exposed had a profound effect not only on the frequency but on the nature of the induced mutations (reviewed in Russell *et al.* 1990, Russell, 2001). One of the chemicals investigated, *N*-ethyl-*N*-nitrosourea (ENU), was found to be by far the most effective mutagen identified in the mouse (Russell *et al.*, 1979b;

Russell *et al.*, 1981). Their work revealed that with ENU, as with radiation, there are DNA repair mechanisms active in mouse germ cells (Russell *et al.*, 1982b; Russell *et al.*, 1982c).

An SLT utilizing a different test stock was conducted by a Medical Research Council group at Harwell, England. This test stock carried homozygous recessives at six specific loci, namely: fuzzy (*fz*) and leaden (*ln*) on chromosome 1; brachypody (*bp*; growth differentiation factor 5, *Gdf5*), nonagouti (*a*), and pallid (*pa*; *Pallidin*, *Pldn*) on chromosome 2; and pearl (*pe*; beta-3A adaptin, *Ap3b1*) on chromosome 13. Because of the selection of a different set of specific loci from those in the Russell T stock, the Harwell data indicated a slightly different average mutation rate from that calculated by the Russells (Lyon and Morris, 1966). However, their conclusions were identical to those drawn from the work at Oak Ridge: ionizing radiation induces heritable mutations in mammals at a higher rate than it does in *Drosophila melanogaster* (Russell, 1951; Lyon and Morris, 1966).

DELETION COMPLEXES IN THE MOUSE GENOME

The radiation mutagenesis work at ORNL led to the isolation of numerous mutations at the seven specific loci. Molecular and genetic analyses of mice carrying these mutations led to the conclusion that many of the mutations are chromosomal deletions of varying sizes. Crosses between pairs of mutants at each locus, followed by a detailed phenotypic analysis of the progeny of these crosses, led to the identification of complementation groups. Subsequent molecular analysis within complementation

groups has allowed each family of deletions to be ordered into a deletion complex, consisting of a set of nested and overlapping deletions (reviewed in Russell and Rinchik, 1987 and Russell, 1989; O'Brien *et al.*, 1996). In addition to those generated at ORNL, deletion complexes have been produced by other groups at some of the same (i.e. the *pink-eyed dilution*) and other loci. These mutations have been induced through mutagenesis schemes utilizing either ionizing radiation in mice (Lyon and Morris, 1966) or embryonic stem (ES) cells (You *et al.*, 1997; Thomas *et al.* 1998; Schimenti *et al.*, 2000), or Cre-LoxP recombination in ES cells (Ramirez-Solis *et al.*, 1995).

Deletion complexes have proven to be an invaluable reagent in the genetic and molecular analysis of the regions of the genome that they cover. Crosses between carriers of different deletions within the same complex are used to identify regions of the genome (called complementation groups) that are associated with a particular phenotype by examining the offspring of such crosses to determine what "functions" are missing. The missing "function(s)" must map to the region of the genome deleted by the overlap of the two deletions (Rinchik and Russell, 1987; Rinchik and Russell, 1990). Such genetic, or complementation, analyses have been used to identify complementation groups in the brown (*b*) region of chromosome 4 (Rinchik, 1994), the albino (*c*) region of chromosome 7 (Gluecksohn-Waelsch, 1979; Russell *et al.*, 1982a ; Russell *et al.*, 1979a; Niswander *et al.*, 1988; Russell, 1979; Russell and Raymer, 1979, Rinchik and Carpenter, 1999), the pink-eyed dilution (*p*) region of chromosome 7 (Lyon *et al.*, 1992; Russell *et al.*, 1995), the dilute-short ear (*d-se*) region of chromosome 9 (Russell, 1971), and the piebald (*s*) region of chromosome 14 (O'Brien *et al.*, 1996).

Perhaps the most valuable use of the deletion complexes is as a reagent for fine-structural mapping by crossing an animal that is either homozygous (carrying two copies of a single allele) or heterozygous (containing two different alleles) for the recessive trait to be mapped to animals carrying chromosomal deletions from the genomic region of interest. If the recessive phenotype maps within the deletion being tested, then the hemizygous offspring (containing a single allele opposite a deletion) exhibit the recessive phenotype. If the recessive phenotype does not map within the deletion being tested, then the offspring are heterozygous and exhibit a wild-type phenotype (reviewed in Rinchik and Russell, 1990). These types of trans-complementation crosses have been used extensively to map numerous recessive phenotypes and genes to various deletions, including the albino (*c*) deletion complex on chromosome 7 (Russell *et al.*, 1982a; Rinchik *et al.*, 1993b; Potter and Rinchik, 1993; Potter *et al.*, 1995; Rinchik and Carpenter, 1993 and 1999), the *p* deletion complex on chromosome 7 (Culiat *et al.*, 1993), the *Krd* deletion on chromosome 19 (Ji *et al.*, 1999), and the Brachyury (*T*) deletion complex on chromosome 17 (Bergstrom *et al.*, 1998).

The deletion complexes can also be used to fine-map molecular markers and probes relative to the deletion breakpoints. This is accomplished through the use of panels of DNAs from F_1 hybrids of an interspecific cross, *e.g. Mus musculus* X *Mus spretus* cross, with the chromosomal deletion of interest carried on the *Mus musculus* chromosome. Taking advantage of the high rate of polymorphism between the *Mus musculus* and related *Mus* species (Silver, 1995), any molecular probe showing a

restriction fragment length polymorphism (RFLP) between two species can be used for mapping. If the probe maps within the deletion being tested, then the hybridization band corresponding to the *Mus spretus* allele, *e.g.*, will be present and the band corresponding to *Mus musculus* (which carries the deletion) will be absent. On the other hand, if the probe maps outside the deletion, then both the *Mus spretus* and the *Mus musculus* bands will be present. The band corresponding to the *Mus spretus* allele will always be present; therefore, the presence or absence of the *Mus musculus* band indicates whether or not the probe is deleted (reviewed in Russell and Rinchik, 1987). Similarly, any PCR amplification product that shows a polymorphism between *Mus musculus* and *Mus spretus* can also be used for mapping, with the presence or absence of the *Mus musculus* product determining whether or not the product maps within the deletion (Metallinos *et al.*, 1994). Fine-structure mapping has been utilized to molecularly characterize the genomic regions surrounding several of the deletion complexes, including the albino (*c*) deletion complex on chromosome 7 (Johnson *et al.*, 1989; Rinchik *et al.*, 1989; Holdener *et al.*, 1995; Rikke *et al.*, 1997), the *p* deletion complex on chromosome 7 (Johnson *et al.*, 1995; Dhar and Johnson, 1997; Culiat *et al.*, 1993; Culiat *et al.*, 1994; Nicholls *et al.*, 1993; Scrable *et al.*, 1990; Stubbs *et al.*, 1994; Walkowicz *et al.*, 1999; Wu *et al.*, 2000), the brown (*b*) deletion complex on chromosome 4 (Rinchik *et al.*, 1994; Bell *et al.*, 1995), the piebald (*s*) deletion complex on chromosome 14 (Metallinos *et al.*, 1994; O'Brien *et al.*, 1996), the Brachyury (*T*) deletion complex on chromosome 17 (You *et al.*, 1997), and the interdigitated Huntington disease (*Hdh*), dipeptydylpeptidase VI (*Dpp6*), and γ-aminobutyric acid

receptor, β-1 subunit (*Gabrb1*) deletion complexes on chromosome 5 (Schimenti *et al.*, 2000) .

One advantage of using the deletion complexes for mapping, both genetic and molecular, is that the resolution is much higher than that typically obtained using traditional recombination mapping techniques (Rinchik and Russell, 1990; Bergstrom *et al.*, 1998; Schimenti *et al.*, 2000). In addition, any probe or amplification product that is polymorphic can be mapped relative to the deletion breakpoints. This means that entire genomic or cDNA clones, as well as anonymous clones and PCR products, can be mapped, even if their genetic or biochemical function is unknown. The maps produced by these types of analyses are dynamic; they continually evolve as the genes and functional units in these regions become better understood.

Another important use of the deletion complexes and other chromosomal rearrangements is in regional mutagenesis studies (Rinchik *et al.*, 1990). In fact, deletions have been called the "critical genetic reagent" (Schimenti and Bucan, 1998) that, along with the supermutagen ENU, has fueled the explosion of regional mutagenesis experiments in the mouse in progress worldwide (Justice *et al.*, 1999). By mutagenizing a region of the genome corresponding to any of the deletion complexes, it is possible to identify mutations through a hemizygous, two-generation breeding scheme, rather than the three-generation screen that would be required to identify them homozygously. Furthermore, any recessive mutations detected in the hemizygous breeding scheme presumed to map within the genomic location defined by the extent of the deletion can be quickly fine-mapped using *trans*-complementation crosses to

deletion carriers (Rinchik and Russell, 1990; Rinchik, 1991; Schimenti and Bucan, 1998; Schimenti *et al.*, 2000).

Deletion complexes have also been utilized in other ways. For instance, the *Krd* deletion has been used as a reagent in a sensitized screen for genes affecting kidney and retinal function. This is possible because the paired box gene 2 (*Pax2*) maps within the *Krd* deletion, and haploinsufficiency of *Pax2* causes kidney and retinal abnormalities (Ji *et al.*, 1999). The deletion complexes have also be used to determine the biological effect(s) of segmental haploidy on the specific regions of the genome covered by these complexes (Schimenti *et al.*, 2000). Mice carrying chromosomal deletions in the *p*-deletion complex of mouse chromosome 7 (Mmu7) have been used to study an obesity phenotype that appears to be imprinted (Dhar *et al.*, 2000), and mice carrying other deletions are being used to dissect contiguous gene syndromes such as Wolf-Hirshhorn syndrome (Schimenti *et al.*, 2000) and DiGeorge syndrome (Lindsay *et al.*, 1999).

Deletions and deletion complexes produced in the mouse are reagents of unparalleled importance to mouse geneticists. The genetic and molecular analysis of genomic regions corresponding to these deletions and deletion complexes has produced a constantly evolving, high-resolution map of these regions. Recessive phenotypes and mutations, as well as cDNA and genomic clones and other molecular probes and markers (even anonymous markers), can be mapped relative to the deletion breakpoints within the complex to produce a higher resolution map than can be produced through traditional linkage analysis. Deletions and deletion complexes can be used to study

genomic imprinting, to dissect the phenotypes associated with contiguous gene syndromes, and as a genetic reagent in regional mutagenesis programs.

THE *pink-eyed dilution* (*p*) LOCUS

The *pink-eyed dilution* (*p*) locus of the mouse is an historically important region of the mouse genome. The first report of genetic linkage in mammals was between *p* and the albino (*c*), or tyrosinase (*tyr*) locus in the mouse (Haldane *et al.*, 1915). In addition, *p* is one of the seven genetic loci selected for use in the mouse SLT (Russell, 1951).

There have been over 100 mutations identified and reported at the *p* locus of mouse chromosome 7 (Lyon *et al.*, 1992; Brilliant *et al.*, 1994; Johnson *et al.*, 1995). At least 45 of these *p* mutations are classified as chromosomal deletions of varying sizes, and most of these deletions are homozygous-lethal (Russell *et al.*, 1995). In the mouse, the original *p* mutation, which is a recessive and hypomorphic allele (Rinchik *et al.*, 1993a), is of *Mus musculus* (Asian) origin, while the wild-type (*P*) allele is of *Mus domesticus* origin (Brilliant *et al.*, 1994). Mutations at the *p* locus are associated with reduced pigmentation of both the eyes and skin, with a reduction in the quantity of eumelanin (black-brown pigment) produced by the melanosomes as well as an alteration of the size and shape of the melanosomes (Russell, 1949). The protein encoded by the *P* gene is hydrophobic, with 12 putative transmembrane domains (Rinchik *et al.*, 1993a). It has a mass of 110 kDa and is localized to the membrane of the melanosome (Rosemblat *et al.*, 1994).

In humans, the *P* locus maps to chromosome 15q11.2-q12, and mutations at or near this locus have been associated with several diseases, including Prader-Willi and Angelman Syndromes and hypomelanosis of Ito (Gardner *et al.*, 1992; Brilliant, 1992). Mutations in the human *P* gene itself have been shown to cause tyrosinase-positive oculocutaneous albinism type II (OCA2) (Ramsay *et al.*, 1992; Rinchik *et al.*, 1993a; Lee *et al.*, 1994). The human P protein contains 25 exons (the first of which is non-coding) and covers a genomic interval of between 250- and 600-kb. The protein is 838 amino acids long and, like the mouse protein, contains 12 putative transmembrane domains and is localized to the melanosome membrane. The protein shows homology to several prokaryotic transport proteins, including those from species such as *Escherichia coli*, *Staphylococcus aureus*, and *Mycobacterium leprae* (Lee *et al.*, 1995). Therefore, although the exact function of the P protein has not been demonstrated, it is hypothesized that this protein is likely involved in the transport of tyrosine or some intermediate in the tyrosine metabolic pathway (Rinchik *et al.*, 1993a).

THE *pink-eyed dilution* (*p*) DELETION COMPLEX

The *pink-eyed dilution* (*p*) deletion complex consists of a series of nested, overlapping chromosomal deletions produced by the radiation mutagenesis programs at the Oak Ridge National Laboratory and at the M.R.C. Radiobiology Unit at Harwell, England. The physical distance covered by the *p*-deletion complex is unknown; however, it must span at least 6.6 cM of genetic distance, since *p* and *Myod1* are 5.5 cM

apart and *Gabrb3* and *D7Cwr15* are 1.1 cM apart, and all are contiguous within the complex (Johnson *et al.*, 1995) (Figure II-1, in the Appendix).

The *p*-deletion complex has been grossly molecularly characterized, and the current genetic and physical map of the region includes gene loci, microsatellite markers, and functional units (such as *ruby-eye 2*; *ru2*) (Russell, 1989; Johnson *et al.*, 1995; Stubbs *et al.*, 1994; Dhar and Johnson, 1997; Wu *et al.*, 2000; Scrable *et al.*, 1990; Walkowicz *et al.*, 1999; Nicholls *et al.*, 1993; Culiat *et al.*, 1993; Culiat *et al.*, 1994). The proximal (nearest the centromere) region of the complex is homologous to human chromosome 11p14-p15 (Scrable *et al.*, 1990; DeBry and Seldin, 1996), while the distal (nearest the telomere) region of the complex, including *P* itself, is homologous to human chromosome 15q11-q13 (Nicholls *et al.*, 1993; DeBry and Seldin, 1996). The *p*-deletion complex covers a genomic interval that spans from lactate dehydrogenase-C (*Ldh3*), serum amyloid A1 (*saa1*) and myogenic differentiation antigen 1 (*Myod1*) at the proximal end of the complex to γ-aminobutyric acid receptor, β-3 subunit (*Gabrb3*) and human papilloma virus E6-associated protein (*Hpve6a*), later renamed ubiquitin-protein ligase E3a (*Ube3a*), at the distal end of the complex (Russell and Rinchik, 1987; Scrable *et al.*, 1990; Johnson *et al.*, 1995). The limits of the complex are defined by the proximal breakpoint of the most proximally extending deletion, $p^{46DFiOD}$, and the distal breakpoint of the most distally extending deletion, p^{30PUb} (Russell, 1989; Johnson *et al.*, 1995) (Figure II-1).

N-ethyl-*N*-nitrosourea (ENU)

N-ethyl-*N*-nitrosourea (ENU) is the most effective mutagen identified in the mouse, with a mutation rate 87 times higher than the spontaneous mutation rate and over ten times the maximum mutation rate of a single dose of x-rays (Russell *et al.*, 1979b; Hitotsumachi *et al.*, 1985). It is an alkylating agent that is both a toxin and a carcinogen (reviewed in Justice *et al.*, 2000). It can induce mutations in somatic and germ cells in a variety of organisms; therefore, it can be used to induce somatic tumors as well as germline mutations (Davis *et al.*, 1999). ENU causes the highest mutation rate in stem cell spermatogonia; however, it has been shown to cause mutations in mature and maturing oocytes and post-spermatogonial stem cells at a lower rate (Russell and Russell, 1992; Russell *et al.*, 1979b). ENU is widely used in mutagenesis experiments because of its high efficiency and its pronounced effects in spermatogonial stem cells; mutagenized males can produce large numbers of offspring carrying these mutations throughout their reproductive lifetime (Rinchik, 1991; Schimenti and Bucan, 1998).

ENU alkylates DNA through an SN_1 (unimolecular nucleophilic substitution) reaction mechanism in which the ethyl group of ENU is transferred to one of several nucleophilic oxygen or nitrogen atoms in any of the four bases of DNA. The sterically altered (ethylated) base causes mispairing during DNA replication, resulting in a point mutation in subsequent generations (reviewed in Noveroske *et al.*, 2000). A compilation of sequence data from 62 ENU-induced alleles representing 24 genomic

loci indicate that ENU modifies A/T base pairs most frequently, with 82% of mutations falling into this category (Justice *et al.*, 1999).

The mutational spectrum of ENU analyzed to date includes missense mutations, nonsense mutations, and splicing errors. These mutations can be translated into hypomorphic, antimorphic, hypermorphic, and amorphic proteins (Noveroske *et al.*, 2000). ENU mutagenesis studies have produced multiple alleles at several loci, including *short ear* (*Bmp5*) (Marker *et al.*, 1997), *myosin-VA* (*MyoVa*) (Huang *et al.*, 1998a; Huang *et al.*, 1998b), *juvenile development and fertility-2* (*jdf2*/*Herc2*) (Walkowicz *et al.*, 1999), and *quaking* (*qk*) (Cox *et al.*, 1999). These allelic series are invaluable reagents in the effort to elucidate the *in vivo* function of genes in a mammalian system (Ji *et al.*, 1999; Schimenti *et al.*, 2000), enabling the determination of structure-functional relationships in proteins caused by any number of point mutations, and allowing the dissection of the biochemical pathways in which these genes function.

In mouse spermatogonia, there is a threshold level below which ENU shows no mutagenic effect, indicating that there are cellular mechanisms in place that can repair the DNA damage done by ENU, provided that the damage is not too severe (Russell *et al.*, 1982b). Therefore, fractionating a moderate dosage of ENU leads to a lower mutation rate (Russell *et al.*, 1982c). Conversely, it has been shown that weekly, repeated doses of ENU increase the mutation rate and allow the animals to survive a higher dosage that would, if administered in a single dose, be lethal (Hitotsumachi *et al.*, 1985). It has also been found that the effects of ENU can vary in the mouse from

strain to strain; therefore, genetic background has an effect on ENU susceptibility (Davis *et al.*, 1999; Justice *et al.*, 2000).

ENU is a potent point mutagen that primarily alters A/T base pairs in DNA. It is currently the most effective mutagen tested in the mammalian germline, and its effects are most pronounced in mouse spermatogonial stem cells. One of the greatest advantages of ENU is its ability to produce an allelic series at the locus of interest. Its high mutation efficiency, coupled with its ability to produce allelic series, make ENU the mutagen of choice for most mutagenesis experiments.

REGIONAL MUTAGENESIS

There are several regions of the mouse genome where deletions have been utilized as genetic reagents for regional mutagenesis studies. These studies involve using ENU to mutagenize a region of the genome that is covered by a deletion, to effectively "saturate" the genomic region with point mutations induced by ENU and recover these mutations hemizygously using long deletions (reviewed in Rinchik and Russell, 1990). It has been suggested that a true "saturation mutagenesis" is probably not possible, however, due to the variable mutability of different genetic loci (Rinchik and Carpenter, 1999). The advantages to this method are numerous. ENU is the most potent mutagen in mouse spermatogonial cells, so it is ideal for mutagenizing male mice with high efficiency. It produces primarily point mutations, so many different mutations are possible for each gene; therefore, it is possible to produce allelic series, which are useful in determining the functions of the genes within the context of the

animal. Finally, the use of deletions allows the mutations to be recovered hemizygously in only two generations, rather than the three generations that would be required to recover them homozygously (Rinchik, 1991).

REGIONAL MUTAGENESIS EFFORTS AT ORNL

There have been three major regional mutagenesis projects performed at ORNL. In each case, the mutagenesis projects involve the use of ENU as the mutagen and deletions produced from the radiation mutagenesis experiments at ORNL to recover the mutations in hemizygotes. Since the deletions are opposite an easily identifiable morphological marker, the test and carrier classes are conveniently "marked" by their morphological features, such as coat and eye color (Rinchik, 2000). The ORNL mutagenesis efforts have utilized five deletions at three loci: one deletion from the albino (*c*; *Tyr*) deletion complex (Rinchik *et al.*, 1990; Rinchik and Carpenter, 1993 and 1999; Rinchik *et al.*, 1993b), two deletions from the brown (*b*; *Tyrp1*) deletion complex (Rinchik, 2000; Dabney Johnson, personal communication), and two deletions from the *pink-eyed dilution* (*p*) deletion complex (Rinchik *et al.*, 1995; Rinchik, 2000).

REGIONAL MUTAGENESIS WORK IN THE *p*-DELETION COMPLEX

The ORNL mutagenesis work in the *p*-deletion complex has used two different deletions, $p^{46DFiOD}$ and p^{30PUb}. Although these deletions are overlapping and both delete the *p* gene, the proximal breakpoint of $p^{46DFiOD}$ and the distal breakpoint of p^{30PUb} define the proximal and distal limits, respectively, of the *p*-deletion complex (Figure II-1).

Both of these deletions are homozygous-lethal (Russell *et al.*, 1995), and the $p^{46DFiOD}$ deletion, which is cytologenetically visible (Rinchik and Russell, 1990), deletes both *p* and *ru2*, another coat and eye color mutation located approximately 3 cM proximal to *p* (Russell, 1989). The p^{30PUb} deletion deletes *p* and the γ_3, α_5, and β_3 subunits of the *γ-aminobutyric acid receptor* (*Gabrg3*, *Gabra5*, and *Gabrb3*) as well as the *ubiquitin-protein ligase E3a* (*Ube3a*) (Johnson *et al.*, 1995).

MUTAGENESIS USING THE $p^{46DFiOD}$ DELETION

The mutagenesis protocol utilized for this project involves the $p^{46DFiOD}$ deletion, which has two morphological markers, *p* and *ru2*. The mutagenized males are homozygous *ru2 p / ru2 p*; therefore, the mutagenized chromosome carries the both the original *p* mutation and *ru2*, another recessive mutation that produces deep ruby-colored eyes and a somewhat diluted gray-brown coat (Silvers, 1979). These males are then crossed to females that are homozygous wild-type (+/+) to produce wild-type G_1 offspring (unless a dominant mutation is induced). The G_1 offspring are then mated to animals that carry the $p^{46DFiOD}$ deletion (represented by Δ) opposite *ru2*. This cross can produce three classes of animals: *pink-eyed*, *ru2* animals (*ru2 p* / Δ), which would be the test class; *ru2* animals (*ru2 p / ru2* +, with + representing the wild-type allele, *P*), which are the carrier class; and wild-type (either *ru2* + / + + or + + / Δ). Induced recessive mutations are detected in the *pink-eyed*, *ru2* test class of animals, and deleterious or lethal mutations are recoverable in the *ru2* carrier class of animals (Rinchik, *et al.*, 1995) (Figure II-2).

CHAPTER III

MAPPING THE *psrt* PHENOTYPE

There are two stocks of *psrt* animals at ORNL, 723SJ and 1060SJ, independently generated in the *p*-region mutagenesis program. These mutations map between the *p* and *ru2* loci on Mmu 7 (Figure III-1). They exhibit identical phenotypes, which were originally called "small nervous," and they are non-complementing. *Trans*-complementation data indicate that both of these mutations map to the same *p*-deletion interval (E. Rinchik and D. Carpenter, unpublished data), suggesting that these mutations are allelic.

Trans-complementation data show that only four of the *p*-deletions do not complement these two *psrt* phenotypes: $p^{46DFiOD}$, p^{47DTD}, $p^{3RD300H}$, and p^{2MNURf}. All other *p*-deletions tested complement the phenotype (E. Rinchik and D. Carpenter, unpublished data), thereby narrowing down the minimal genomic region known to contain the mutant phenotype to a region of Mmu 7 bordered proximally by the proximal breakpoint of p^{2MNURf} and bordered distally by the proximal breakpoints of p^{8OK} and p^{8R250M}, which cannot be resolved at this time (Figure III-1). The $p^{3RD300H}$ deletion does not complement the *psrt* phenotype; however, at the beginning of this project, its breakpoints had not been well defined and were better characterized in the course of this study.

CHAPTER IV

PHYSICAL MAPPING IN THE *psrt* REGION OF MOUSE CHROMOSOME 7

ABSTRACT

A physical map was produced in order to positionally clone the gene responsible for *psrt*, a profound seizure and runting phenotype that maps into the *p*-deletion complex of Mmu 7. The minimal genomic region determined to contain *psrt* has been further defined by molecularly mapping the breakpoints of the $p^{3RD300H}$ deletion. Two bacterial artificial chromosome (BAC) libraries were screened using molecular markers from the region, and the BACs identified in the screen were further characterized to determine their size and to generate sequence tagged sites (STSs) from their terminal, as well as internal, sequences. BAC STSs and other molecular markers were mapped using deletion panel blots and Southern blots of restriction digested BAC DNAs, and data produced were used to assemble and order the physical map. A physical map of a portion of the *psrt* minimal genomic region has been produced, and BAC clones from this map were utilized (see Chapter V) to identify transcriptional units in an attempt to identify the gene responsible for the *psrt* phenotype.

INTRODUCTION

1. Positional Cloning

This project used positional cloning to identify the gene responsible for the *psrt* phenotype in the mouse. Positional cloning is a technique that was first utilized in the

1980s to clone the human gene responsible for chronic granulomatous disease (Royer-Pokora *et al.*, 1986). Since that time, positional cloning has been used to clone a number of genes in a variety of organisms, including the human genes responsible for Duchenne muscular dystrophy, cystic fibrosis, type 1 neurofibromatosis, fragile X syndrome, and numerous other diseases (Collins, 1992; Talbot and Schier, 1999; Boehm, 1998). The aim of positional cloning is to isolate and clone a gene based solely on its chromosomal position in the absence of information about the nature of function of either the gene or its protein product (Collins, 1992).

Positional cloning projects have three basic steps: identifying and mapping linked markers, constructing a physical map of the genomic interval determined to contain the gene, and searching for genes within this region (Stubbs, 1992; Collins, 1992). In the first step, identifying and mapping linked markers, pedigrees in which the gene of interest is segregating are collected and analyzed (Collins, 1992). This can be done either in humans or in model organisms, such as the mouse, and entails analyzing polymorphic markers from the region to determine evidence of linkage between the mutant phenotype and any of these markers (Collins, 1992; Boehm, 1998). By determining which markers are genetically linked to the mutation, it is possible to establish the limits for the region of the genome containing the mutation. The second step is the construction of a physical map, consisting of a contiguous, overlapping series of genomic clones covering the region of the genome identified in step one (Stubbs, 1992). The third step in a positional cloning strategy is identifying transcriptional units

within the genomic interval of interest and analyzing each of these for suitablilty as a candidate for the mutant phenotype (Collins, 1992).

2. Positional Candidate Approach

The positional candidate approach, which is a separate but related strategy, has been successfully utilized to clone several genes, including those responsible for amyotrophic lateral sclerosis (Lou Gehrig's disease), Waardenburg syndrome, and Charcot-Marie-Tooth disease (Ballabio, 1993). The positional candidate approach can be used when some biochemical or functional information is known about the disease. The disease phenotype is mapped in a similar manner to a positional cloning strategy; however, it is not necessary to go through the laborious process of producing a physical map of the region. Once the disease has been mapped, then the available databases are searched for genes and expressed sequence tags (ESTs) mapping into the region, and candidate genes are proposed by evaluating likely functional suitability and by looking for mutations in their transcripts (Ballabio, 1993).

3. Mapping Markers in the *psrt* Region of Mouse Chromosome 7

The first step in any positional cloning strategy is to identify and map markers linked to the mutation to define the limits of the minimal genomic region containing the mutation. Any markers identified and mapped in this step can be used to "anchor" the physical map (Stubbs, 1992; Collins, 1992). In the mouse, this is readily accomplished through the use of polymorphic markers, such as microsatellites (Boehm, 1998).

A *trans*-complementation analysis of the *psrt* mutation has narrowed down the minimal genomic region known to contain the mutation to a region of Mmu 7 bordered proximally by the proximal breakpoint of p^{2MNURf} and bordered distally by the proximal breakpoints of p^{8OK} and p^{8R250M}, which cannot be resolved at this time (Figure III-1). The $p^{3RD300H}$ deletion does not complement the *psrt* phenotype; however, its breakpoints have not been well defined. An earlier molecular analysis of 36 mutations at the *p* locus determined that the $p^{3RD300H}$ deletes *growth arrest 2* (*Gas2*), *D7H15F37S1*, and *p*, three molecular markers, as well as *l7Rl1*, a phenotype-defined marker that maps between *Gas2* and *D7H15F37S1*. The proximal breakpoint of the $p^{3RD300H}$ lies in the genomic region between *Gas2* and *ru2*, and the distal breakpoint lies between *p* and *Gabrg3* (Johnson *et al.*, 1995). A subsequent molecular analysis of the *p*-deletion complex mapped fourteen microsatellite markers relative to the previously defined deletion breakpoints (Dhar and Johnson, 1997). The microsatellite study, however, included only 23 of the *p*-region deletions; so not all of the deletions utilized for the first study were included in the second study. Since the $p^{3RD300H}$ was one of the *p*-deletions that did not complement the *psrt* phenotype and was not included in the microsatellite study, further defining the proximal breakpoint of this deletion may serve to narrow down the minimal genomic region known to contain the *psrt* mutation.

4. Constructing the Physical Map of the *psrt* Region

Constructing a physical map of the minimal genomic interval determined to contain the gene is the second step in the process of positionally cloning the gene

responsible for the *psrt* phenotype. Generating and fine-mapping new markers in the region relative to the deletion breakpoints and to the BAC clones in the region will allow the clones to be assembled into a contig covering the region. The clones from the contig will then be analyzed to identify transcriptional units, and an expression analysis of the *psrt* and control animals will be used to assess the candidacy of any genes identified in the process.

MATERIALS AND METHODS

1. Deletion Panel Mapping of Microsatellites

D7Mit230 and *D7Mit312*, two microsatellite markers that show size polymorphism between *Mus musculus* and *Mus spretus*, were mapped by PCR using genomic DNAs from interspecific hybrid (*Mus musculus* X *Mus spretus*)F_1 carriers of various *p* deletions in which a different *Mus musculus p*-region deletion chromosome was balanced opposite a *Mus spretus* chromosome (Johnson *et al.*, 1995; Dhar and Johnson, 1997). F_1 interspecific hybrids were generated as described previously (Johnson *et al.*, 1989).

Oligonucleotide primers for the microsatellite markers were obtained from Research Genetics Inc. (Huntsville, AL), and PCRs were performed using a 20 μL total reaction volume containing 10-50 ng of template DNA, 2 Units of Taq DNA Polymerase (Gibco BRL Life Technologies, Inc.; Gaithersburg, MD), 0.6 μM each primer (forward and reverse), and 14 μL of PCR mix to give a final concentration of 1.5 mM $MgCl_2$ and 1X PCR buffer. Reactions were carried out in 0.2 mL PCR tubes using

a PTC-100 thermocycler (MJ Research, Waltham, MA) and the following program: initial denaturation at 94°C for 5 minutes followed by 30 cycles of denaturation at 94°C for 45 seconds, annealing at 45°C for 45 seconds, and extension at 72°C for 150 seconds, then a final five minute extension step at 72°C. PCR products were analyzed on 3% NuSieve GTG agarose gels (FMC Bioproducts, Rockland, ME) made with 1X TAE (0.04 M Tris-acetate, 0.001 M EDTA) and visualized under UV light after staining with ethidium bromide.

2. Defining the Breakpoints of $p^{3RD300H}$

The breakpoints of $p^{3RD300H}$ were defined by PCR amplification of eight microsatellite markers using genomic DNA from interspecific hybrid (*Mus musculus* X *Mus spretus*)F_1 carriers of the $p^{3RD300H}$ deletion. The eight microsatellite markers used were *D7Mit26*, *D7Mit69*, *D7Mit70*, *D7Mit84*, *D7Mit145*, *D7Mit160*, *D7Mit193*, and *D7Mit230*. Reactions were performed, and PCR products were analyzed, as described above.

3. BAC Library Screenings

The Genome Systems, Inc. (St. Louis, MO) "Down to the well" Mouse (strain 129SV) ES BAC DNA pools were screened by PCR using primers for *D7Mit145*, *D7Mit193*, *D7Mit194*, and *D7Mit230*. The reactions were performed using 3 μL of the BAC superpool DNA in a 20 μL total reaction volume, using the same reagents and

thermocycling program as described above, annealing at 45°C (or 48°C for *D7Mit145*) for 35 cycles. PCR products were analyzed as described above.

The Genome Systems, Inc. (St. Louis, MO) BAC Mouse Library (Release II) high density filters were also screened by hybridization using RT 591/592, a 287 bp RT-PCR product from the mouse *Prmt3* cDNA. The generation and analysis of this RT-PCR product is described in detail in Chapter VI (Sequence and Expression Analyses of Candidate Genes). Standard hybridization conditions were utilized, and blots were washed three times, with a final 20 minute wash at 65°C in 0.1 X SSCP. BACs screened by hybridization were confirmed by PCR, using primer pair ORN 591/592, annealing at 50°C for 35 cycles. BACs were grown in LB medium plus chloramphenicol (12.5 μg/mL) and BAC DNA was isolated using either the Promega Wizard® Plus Minipreps DNA Purification System (Promega Corporation, Madison, WI) or the Qiagen Plasmid Midi Kit (Qiagen Inc., Valencia, CA) according to the manufacturer's instructions.

4. Determining the Size of BAC Inserts

BACs were grown and the DNA was isolated as described above. BAC DNAs were digested overnight at 37°C with *Not*I (Gibco BRL Life Technologies, Rockville, MD) using the following reaction cocktail: 10μL of miniprep or midiprep DNA, 2.5 μL of 10 X React3 and 2 μL (15 units/μL) *Not*I (Gibco BRL Life Technologies, Rockville, MD), and 10.5 μL of sterile water. Digested DNAs were loaded on a 1% agarose gel which was run in 0.5X TBE at 14°C. Electrophoresis was performed in a BioRad Chef

Mapper pulsed field gel electrophoresis (PFGE) apparatus using the autoalgorithm function programmed to separate fragments ranging from 50-kb to 500-kb. The gel was stained with ethidium bromide and visualized under UV light. The size of the inserts was determined by comparison to low-range and mid-range PFGE markers (New England Biolabs, Beverly, MA).

5. BAC Endcloning and STS Generation

STSs were generated from the terminal sequences of the BACs using a modification of the ligation linker PCR method (Riley *et al.*, 1990; Mueller and Wold, 1989). Briefly, the BAC is digested with several different restriction enzymes, an oligo linker is ligated to the fragments, and PCR is performed on the digested fragments using primers for the oligo linker and either the SP6 or T7 promoter sequences, which are in the BAC vector near the cloning site (Shizuya *et al.*, 1992; Kim *et al.*, 1996). STSs were also generated from internal sequences by single primer PCR using primers for the B1 and B2 repeat sequences in the mouse (Kim *et al.*, 1997a).

The PCR products obtained were cloned using the pGEMT-Easy II cloning kit (Promega Corporation, Madison, WI) and analyzed by fluorescent sequencing with the Big Dye Terminator Kit (PE Biosystems, Foster City, CA). The products were run on an ABI377 DNA sequencer and the sequence analyzed using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) to verify the identity of the cloned products. STS clones were verified by mapping back to the *p*-region of Mmu 7 using deletion

panel Southern blots and standard hybridization protocols (Johnson *et al.*, 1995; Dhar and Johnson, 1997).

6. Placing STSs on the Physical Map

STSs were placed on the physical map in one of two ways: by PCR or by hybridization. Nine microsatellite markers (*D7Mit26*, *D7Mit69*, *D7Mit70*, *D7Mit84*, *D7Mit145*, *D7Mit160*, *D7Mit193*, *D7Mit229* and *D7Mit230*) and one EST (RT 591/592) were amplified by PCR using BAC miniprep DNAs and conditions described above (D7Mit229 was annealed at 40°C for 35 cycles). *D7Mit229* was used with the BAC DNAs, but not for mapping with the deletion panels, because it is not polymorphic between *Mus musculus* and *Mus spretus*. BACs positive for the STS will show an amplification product of the expected size.

Other STSs were placed on the physical map by hybridizing to Southern blots of BAC DNAs digested with *EcoR*V. BACs positive for the probe show a hybridization band of the appropriate size in the autoradiograph of the washed blot. Seven STSs were placed on the physical map by hybridization, including four BAC end clones (B36n10/EcoRV/SP6, B36n10/EcoRV/T7, B179d14/EcoRV/SP6, and B179d14/EcoRV/T7), and two RT-PCR products from the *Prmt3* gene, RT 595/596 and RT 591/592, which was also mapped by PCR (Table IV-1, in the Appendix).

7. Deletion Panel Mapping of Probes

STS clones and other probes showing a restriction fragment length polymorphism (RFLP) between *Mus Musculus* and *Mus spretus* were verified by mapping to the *p*-deletion complex by hybridization using deletion panel Southern blots and standard hybridization protocols (Johnson *et al.*, 1995; Dhar and Johnson, 1997). Probes were isolated on 1% LMP (low melt point) Agarose (Gibco BRL, Rockville, MD) then melted and random prime labeled with [α-^{32}P]dCTP. Hybridizations were performed at 42°C overnight, and blots were washed three times: 20 minutes at room temperature in 2X SSCP with 0.1 % SDS, followed by 20 minutes at 65°C in 2X SSCP and 20 minutes at 65°C in 0.1X SSCP. The following probes were mapped by hybridization to deletion panel blots: B36n10/B2A, B4514/Pvu/T7, B179d14/B1B, B179d14/EcoRV/SP6, B179d14/EcoRV/T7, RT 591/592.

RESULTS

1. Deletion Panel Mapping of Microsatellites

In order to better define the region of the genome surrounding the *psrt* locus, two microsatellite markers were mapped. A PCR analysis of *D7Mit230* using the deletion panel DNAs shows that the *Mus musculus* band is absent for $p^{46DFiOD}$, p^{47DTD}, p^{2MNURf}, and $p^{3RD300H}$, while it is present in p^{8OK} and p^{8R250M}. The absence of the *Mus musculus* band indicates that the marker is deleted; therefore, the marker must lie between the proximal breakpoints of p^{2MNURf} and $p^{3RD300H}$ (which are unresolved at this time) and p^{8OK} and p^{8R250M} (which are also unresolved). A similar analysis of *D7Mit312*

shows that the *Mus musculus* band is present for all six deletion DNAs, which means that the marker is not contained within the deleted region for any of the six deletions tested. Therefore, *D7Mit312* must map outside the *p*-deletion complex (Figure IV-2).

2. Defining the Breakpoints of $p^{3RD300H}$

PCR analysis of eight microsatellite markers in the *p*-deletion complex has further defined the breakpoints of $p^{3RD300H}$. Three of the microsatellite markers tested were deleted in $p^{3RD300H}$: *D7Mit193*, *D7Mit230*, and *D7Mit145*. Five of the microsatellite markers were not deleted: *D7Mit26*, *D7Mit69*, *D7Mit70*, *D7Mit84*, and *D7Mit160*. This means that the proximal breakpoint of $p^{3RD300H}$ is between *D7Mit26* and *D7Mit69* (which cannot be resolved at this time) and *D7Mit193* and *D7Mit230* (which also cannot be resolved). Previously, the proximal breakpoint of this deletion was known to be somewhere between *Ldh3* and *Gas2* (Johnson *et al.*, 1995). The distal breakpoint of the $p^{3RD300H}$ deletion is between *D7Cwr15* and *D7Mit84*. Previous mapping studies indicate that the $p^{3RD300H}$ deletion encompasses the *D7H15F37S1* locus (Johnson *et al.*, 1995) and the *D7Mit70* locus maps proximal to *D7H15F37S1*, between *D7H15F37S1* and *D7Mit196* (Wu *et al.*, 2000). These data indicate that *D7Mit70* is not deleted in $p^{3RD300H}$, meaning that $p^{3RD300H}$ is a "skipper," which is a noncontiguous deletion most likely involving some sort of complex rearrangement (Johnson *et al.*, 1995) (Figure IV-3).

3. BAC Library Screenings

A total of eleven BACs from the Genome Systems, Inc. (St. Louis, MO) "Down to the well" Mouse ES BAC DNA pools were screened by PCR using primers for four different microsatellite markers mapping near the *psrt* region. There were four BACs selected using microsatellites mapping within the *psrt* minimal interval: three positive for *D7Mit193* (B36n10, B45l4, and B179d14) and one positive for *D7Mit230* (B211i3). In addition, *D7Mit145*, the microsatellite marker mapping in the distally adjacent deletion interval, was used to identify two BACs (B207p17 and B212f2). *D7Mit194*, which maps in the most proximal deletion interval and may be useful in cloning *ru2*, was used to select five BACs (B47g18, B47g22, B142k10, B142k22, and B178n12).

There were also twelve BACs from the Genome Systems, Inc. (St. Louis, MO) BAC Mouse Library (Release II) high density filters that were identified by hybridization using RT 591/592, a 287 bp RT-PCR product from the mouse *Prmt3* cDNA, as the probe. They are: B242g12, B279l14, B288k21, B315k22, B329d6, B333e23, B344l6, B354k11, B372f7, B382o13, B434g10, and B462a15. These BACs were confirmed by PCR using four different sets of primers: RT 591/592, which was used to amplify the RT-PCR product that was used to screen the library; *D7Mit193*, which maps in the *psrt* minimal interval; ORN 609/610, which corresponds to the T7 end of B45l4; and ORN 607/608, which corresponds to the terminal sequence of B179d14. The BACs were then subdivided into categories based on the results of the PCRs (Table IV-4).

4. Deletion Panel Mapping of STSs

Deletion panel Southern blots were used to map probes showing an RFLP between *Mus musculus* and *Mus spretus*. Seven probes were mapped this way, including three BAC end clones (B36n10/EcoRV/SP6, B179d14/EcoRV/SP6, and B179d14/EcoRV/T7), and two STSs derived from single-primer PCR (B36n10/B2A and B179d14/B1B). In every instance, the *Mus musculus* bands are present in p^{8OK} and p^{8R250M} and absent in $p^{46DFiOD}$, p^{47DTD}, $p^{3RD300H}$ and p^{2MNURf}, which means that the probes are deleted in in $p^{46DFiOD}$, p^{47DTD}, $p^{3RD300H}$ and p^{2MNURf}, but not in in p^{8OK} and p^{8R250M}. Therefore, these seven probes must map proximal to the p^{8OK} and p^{8R250M} proximal breakpoints and distal to the $p^{3RD300H}$ and p^{2MNURf} proximal breakpoints, which is the genomic region identified as the *psrt* minimal deletion interval (Figure IV-5).

There were two STSs that appear to contain repetitive sequences because they showed smears rather than distinct hybridization bands. Therefore, these STSs could not be used as probes for hybridization and no mapping results could be obtained for them. They are: B45l4/Rsa/SP6 and B211i3/EcoRV/T7a.

5. BAC Insert Sizes

PFGE was used to determine the insert size of six of the BACs used to produce the physical map. The six BACs and their calculated insert sizes are: B36n10 (120-kb), B45l4 (120-kb), B179d14 (150-kb), B207p17 (280-kb), B211i3 (125-kb), B212f2 (190-kb). Four of the BACs have a single internal *Not*I restriction site (B36n10, B45l4, B179d14, and B211i3), one BAC has two sites (B207p17), and one has zero (B212f2)

(Table IV-6). All of the BACs whose insert size was determined were from the Genome Systems, Inc. (St. Louis, MO) "Down to the well" Mouse ES BAC DNA pools.

6. Assembly of the Physical Map

The physical map was assembled by identifying the markers present on each of the clones, checking for overlap of the markers on the BACs, then ordering them into a contig. The map consists of thirteen BACs, seven from the "Down to the well" Mouse ES BAC DNA pools, and six from the BAC Mouse Library (Release II) high density filters. The insert sizes of the BACs from the latter were not determined. Two sets of BACs were determined to contain the same markers: B315k22, B372f7 and B382o13 all contain RT 591/592 and ORN 607/608; and B242g12, B288k21, B344l6, and B462a15 all contain D7Mit193, ORN 609/610, and ORN 607/608. To minimize redundancies, only one of the BACs from each of these two groups is included in the map (Figure IV-7).

DISCUSSION

Fine-mapping the endpoints of the $p^{3RD300H}$ deletion places the proximal breakpoint between the same two sets of microsatellite markers as the p^{2MNURf} deletion. Since both $p^{3RD300H}$ and p^{2MNURf} fail to complement the *psrt* phenotype, their proximal endpoints (which cannot be resolved at this time) must define the proximal limit of the minimal genomic interval determined to contain the *psrt* mutations. The distal limit of

the interval is defined by the proximal breakpoints of p^{8OK} and p^{8R250M}. Therefore, the physical map should cover the entire minimal region, from the proximal breakpoints of $p^{3RD300H}$ and p^{2MNURf} to the proximal breakpoints of p^{8OK} and p^{8R250M}, to ensure that the *psrt* transcriptional unit is included.

The current physical map of the region does not completely span the *psrt* minimal region because there is not uninterrupted coverage from one end of the minimal region to the other (Figure IV-7). The physical size of the region is unknown; therefore, it is impossible to know how complete the map is at this time. The map does, however, contiguously span an area within the *psrt* minimal genomic region that includes three microsatellite markers and ten molecular probes (Figure IV-7). The physical map, while not complete, provides a number of BAC clones that can be analyzed to identify transcriptional units in the region that may be responsible for the *psrt* phenotype.

CHAPTER V

SEQUENCE ANALYSIS OF BAC B179d14 AND IDENTIFICATION OF CANDIDATE GENES

ABSTRACT

Sequence analysis of BAC B179d14 has identified two candidate genes for the *psrt* phenotype: *Tat-interacting protein (30 kD)* (*Tip30*) and *protein arginine N-methyltransferase* (*Prmt3*). A third candidate gene, *glycine transporter type 2* (*Glyt2*), was identified because of its location in the human genomic region of homology and its function in neurotransmission. The sequence of BAC B179d14 was used to determine the structure of a 150-kb genomic interval containing the mouse *Tip30/Cc3* and *Prmt3* genes.

INTRODUCTION

The process of identifying candidate genes is an important part of any positional cloning or positional candidate strategy. These two strategies are related and can be complementary. The positional cloning approach differs from the positional candidate approach in that positional cloning involves the laborious step of producing and assembling a physical map of the genomic region determined to contain the mutation, followed by a search for transcriptional units within the region (Collins, 1992). The positional candidate approach, on the other hand, can be utilized when some biochemical or functional information is known about the disease. Transcriptional

units are identified through searching the available databases for genes and ESTs mapping into the region (Ballabio, 1993). Once candidate genes have been identified using either approach, expression and sequence analyses are required to evaluate the validity of their candidacy by determining if mutations exist in the transcripts (Collins, 1992; Ballabio, 1993). Since some information is known about the *psrt* phenotype, it is possible to utilize both the positional cloning and positional candidate approaches to identify and clone the gene responsible for *psrt*.

Once the physical map is assembled, it is necessary to isolate and identify transcriptional units in the genomic (BAC) clones that compose the physical map. There are numerous techniques available for this purpose (Parimoo *et al.*, 1995). They range from splicing signal-based methods such as exon trapping (Duyk *et al.*, 1990; Krizman and Berget, 1993) and exon amplification (Buckler *et al.*, 1991; Church *et al.*, 1994) to methods relying on the identification of either CpG islands or phylogenetically conserved sequences (Parimoo *et al.*, 1995) to cDNA selection (Parimoo *et al.*, 1991) and direct selection (Lovett *et al.*, 1991) to sequence and computational-based methods utilizing DNA sequencing (Claverie, 1994) and analysis of the sequence utilizing programs such as the gene recognition and analysis internet link (GRAIL) (Uberbacher and Mural, 1991).

The method chosen for this project was random shotgun DNA sequencing followed by computer analysis of the sequence to detect transcriptional units. Sequencing was done by randomly fragmenting the original BAC clone and subcloning the fragments into a vector that is easily sequenced. A comparison of the sequencing

results will identify overlap, which can be used to assemble the sequences into contigs. Contigs are screened for repetitive elements and the sequence is analyzed for transcriptional units (Claverie, 1994).

This project was accomplished using a combination of the positional cloning strategy and the positional candidate approach. A search of the databases occurred concomitantly with the development and assembly of the physical map in an effort to identify the gene responsible for the *psrt* phenotype.

MATERIALS AND METHODS

1. Selecting a BAC for Sequencing

The BAC clone chosen for sequencing was B179d14, from the Genome Systems, Inc. (St. Louis, MO) "Down to the well" Mouse ES BAC DNA pools. B179d14 was one of the first group of BACs screened, and was selected using *D7Mit193*, which maps in the same deletion interval as the *psrt* phenotype. Its size, determined by PFGE, is approximately 150-kb. There are seven STSs contained on B179d14 that have been fine-mapped back to the *psrt* minimal region, and three other STSs that could not be mapped to the deletions (because they do not show an RFLP between *Mus musculus* and *Mus spretus*) but were mapped to BAC clones in the region. This means that B179d14 is most likely not chimeric, because it contains ten molecular markers within 150-kb that map to the same deletion interval.

2. BAC DNA Purification, Subcloning, and Sequencing

BAC DNA was isolated using the Qiagen Mega Prep Kit (Qiagen Inc., Valencia, CA) according to the manufacturer's instructions, except using double the recommended volume for solutions P1, P2, and P3. The BAC DNA was further purified through two cycles of CsCl gradient centrifugation (Sambrook *et al.*, 1989). Approximately 20% of the DNA was then sheared using a Gene Machine Hydroshear on a setting of eight, and the sheared DNA was treated with T4 Polymerase to make blunt ends. The sheared, blunt ended DNA was ligated into the pBSK- vector (Stratagene Inc., La Jolla, CA) that had been *EcoR*V digested and phosphatased. The ligated DNA was electroporated into *E. coli* DH5α cells (Gibco BRL Life Technologies, Rockville, MD) and putative transformants were selected on LB plus Ampicillin (100 μg/mL) plates.

Plasmid DNA from the subclones was isolated on a Qiagen BioRobot 9600 using the QIAprep 96 Turbo BioRobot Kit (Qiagen Inc., Valencia, CA). The templates were then sequenced using either the Dye Primer or Big Dye Terminator Kits (PE Biosystems, Foster City, CA) and the products run on an ABI377 DNA sequencer.

3. Analysis of Sequence Data

Sequence data were processed and assembled using the Phred, Phrap, and Consed package of progams on an SGI 02 workstation. Approximately 2,440 subclones were sequenced to produce approximately 8X coverage. Repeat sequences were identified using RepeatMasker (University of Washington, Seattle, WA). Additional

analyses were performed using either the Wisconsin Package of programs, Version 10 (Genetics Computer Group, Madison WI), the Genome Channel software (ORNL, Oak Ridge, TN; http://www.genome.ornl.gov) or the NCI BLAST program (http://www.ncbi.nlm.nih.gov/BLAST).

4. Database Search for Positional Candidates

The human genome sequencing database at the National Center for Biotechnology Information (NCBI) was searched for candidate genes mapping within the (human) 11p14-p15 genomic interval, which is homologous to the *psrt* region of Mmu 7. The database is located at the NCBI website (http://www.ncbi.nlm.nih.gov/genome/seq). Genes mapping in the 11p14-p15 interval were accessed by following the "human genome resources" link, selecting chromosome 11, and scrolling to the 11p14-p15 interval and zooming in or out, as necessary. There are links for every gene listed which will open a new page with information and links on the gene symbol, locus information, map information, NCBI reference sequence, GenBank sequences, and annotations.

RESULTS

1. Candidate Gene Identification

There were two genes identified from a BLAST search of the repeat-masked sequence data from BAC B179d14: the human Tat-interacting protein (30kD), *TIP30* (Genbank accession # NM_006410), and the *Rattus norvegicus* protein arginine *N*-

methyltransferase 3, *Prmt3* (Genbank accession # AF059530). A search of the human genome sequencing database identified several genes in the human 11p14-p15 interval. Although no known human disease genes were found to map within this interval, the *glycine transporter type 2* (*GLYT2*) gene (Genbank accession # AF085412) was identified as a positional candidate because of its function in neurotransmission.

2. Structure of the Insert of BAC B179d14

The insert of BAC B179d14 is approximately 150-kb in length, as determined by both PFGE and sequencing data. It contains *Tip30/Cc3* and *Prmt3*, both of which are transcribed in the same direction, separated by approximately 5-kb of genomic sequence. The *Tip30/Cc3* gene is approximately 15-kb in length, while the *Prmt3* gene is approximately 85-kb. There is a single *Not*I restriction site in the CpG island near the 5' end of *Prmt3*, and *D7Mit193* is located just 5' of *Tip30/Cc3*. There is approximately 25-kb of sequence 5' of *Tip30/Cc3*, and approximately 20-kb of sequence 3' of *Prmt3* (Figure V-1).

3. Genomic Structure of the *Tip30/Cc3* Gene

The murine *Tip30/Cc3* gene is composed of five exons spanning approximately 15-kb of genomic sequence. The exons range in size from 62 to over 200 bp in size, and the introns range from 660 bp to approximately 8-kb (Table V-2).

4. Genomic Structure of the *Prmt3* Gene

The murine *Prmt3* gene is composed of 16 exons spanning approximately 85-kb of genomic sequence. The exons range from 51 bp to 188 bp in size, and the introns range from 80 bp to over 19-kb (Table V-3).

DISCUSSION

BAC B179d14 was chosen for sequencing for several reasons. It was one of the first BACs screened; as such, it has been well characterized. Both ends of the BAC, as well as an internal STS, have been cloned and fine-mapped to the *p*-deletion complex. It is a fairly large clone, with an insert size of approximately 150-kb. And it is most likely not chimeric, because it contains nine molecular markers within its 150-kb insert that map to the same deletion interval.

The genomic sequence of *Tip30/Cc3* has not been published; however, a similar cDNA, named *Tc3* (for truncated *Cc3*), has been described (Genbank accession # AF092095). The TC3 and CC3 proteins share a similar N-terminal domain; however, TC3 has a unique, shorter C-terminus. While Cc3 is a tumor suppressor that functions through activation of the apoptotic pathway, Tc3 is a tumor promoter that strongly inhibits apoptosis. It has been postulated that *Cc3* and *Tc3* are alternatively spliced transcripts from the same genomic locus (Whitman *et al.*, 2000). Data from this project, however, dispute this claim. A full 110-kb of genomic sequence downstream of *Tip30/Cc3* has been sequenced in this BAC, and the unique portion of the *Tc3* transcript is not present in this sequence in mouse. This 110-kb of sequence includes the 5-kb

intergenic space between *Tip30/Cc3* and *Prmt3*, the 85-kb that comprises the *Prmt3* gene, and 20-kb of sequence downstream of *Prmt3*; therefore, it is unlikely that *Tip30/Cc3* and *Tc3* are alternatively spliced transcripts arising from the same genomic locus.

Sequence analysis of BAC B179d14 has provided two candidate genes for the *psrt* phenotype, *Tip30/Cc3* and *Prmt3*. A search of the human genome sequencing database in 11p14-15, the region of homology, has provided one additional candidate gene: *GLYT2*. Two of these genes have been evaluated for their suitability as candidates for the *psrt* phenotype. All three of these candidates are described in detail in Chapter VI (Description of Candidate Genes).

CHAPTER VI

DESCRIPTION OF CANDIDATE GENES

INTRODUCTION

There are three candidate genes that have been identified for the *psrt* phenotype: *Tat-interacting protein 30 kDA* (*Tip30*), *glycine transporter type 2* (*Glyt2*), and *protein arginine N-methyltransferase 3* (*Prmt3*). What follows is a description of each of the three candidates, and a discussion of the function of the proteins encoded by these genes.

Tat-interacting protein 30 kDa (*TIP30*)

Tat-interacting protein 30 kDa (TIP30) is a cellular protein that interacts with the trans-activator of transcription (TAT) protein, a transcription factor encoded by HIV-1 (Xiao *et al.*, 1998). TAT promotes transcriptional elongation by RNA polymerase II (RNAP II) and requires multiple cofactors, including TIP30, CDK 9, cyclin T, Tat-SF1, P-TEFb, TFIIH, TAK, (human) SPT5, and CA150 (Herrman and Rice, 1995; Zhou and Sharp, 1995; Cujec *et al.*, 1997; Keen *et al.*, 1997; Sune *et al.*, 1997; Xiao *et al.*, 1998; Yankulov and Bentley, 1998; Zhou *et al.*, 1998; Jeang *et al.*, 1999; Kim *et al.*, 1999). Tat is unique in that it binds to an RNA, rather than a DNA, ligand during transcription (Zhou and Sharp, 1995; Parada and Roeder, 1996). Tat binds to TAR, the trans-activation response region, which is a stem-loop structure near the initiation site of the HIV-I transcript (Zhou and Sharp, 1995; Yankulov and Bentley,

1998). Tat has been shown to have other functions as well, including an involvement in HIV-1 reverse transcription (Harrich *et al.*, 1997).

TIP30 is a member of the short-chain dehydrogenase/reductase (SDR) family (Baker, 1999). It contains a $\beta\alpha\beta$ fold near the amino terminus as well as other structural features conserved among members of the SDR family (Baker *et al.*, 2000). TIP30 is the same as the protein called CC3, a known suppressor of metastasis (Liu *et al.*, 1999; Shtivelman, 1997). Human *CC3* is ubiquitously expressed in tissues as well as in various cell lines, including the classic small cell lung carcinoma (c-SCLC) cell line. However, no *CC3* expression has been demonstrated in variant small cell lung carcinoma (v-SCLC) cell lines, which exhibit very aggressive metastatic behavior. It has been hypothesized that CC3 functions as a tumor suppressor by promoting apoptosis (Shtivelman, 1997). TIP30/CC3 has recently been identified as a serine/threonine kinase, and is involved in the Tat-dependent phosphorylation of the C-terminal domain of the largest RNAP II subunit. This kinase activity is necessary for TIP30/CC3 to function both as a transcriptional cofactor and in regulating apoptosis (Xiao *et al.*, 2000).

The genomic sequence of *Tip30/Cc3* has not been published; however, a similar cDNA, named *TC3* (for truncated *CC3*), has been described in human. The TC3 and CC3 proteins share a similar N-terminal domain; however, TC3 has a unique, shorter C-terminus. While CC3 is a tumor suppressor that functions through activation of the apoptotic pathway, TC3 is a tumor promoter that strongly inhibits apoptosis. The TC3 protein is highly unstable and, therefore, short lived. It has been postulated that *CC3*

and *TC3* arise from alternatively spliced transcripts from the same genomic locus since a large portion of their N-terminal domains are identical (Whitman *et al.*, 2000); however, sequencing data from this project dispute this claim.

Glycine transporter type 2 (*GLYT2*)

Neurotransmitter transporters are membrane-bound proteins located either in the plasma membrane or in the synaptic vesicle membrane. There are three types of these transporters: the Na^+/Cl^- and K^+/Cl^- dependent transporters located in the plasma membrane, and the pH or H^+-dependent transporters located in the membrane of the synaptic vesicle (Lesch *et al.*, 1996). Neurotransmitter transporters establish and mediate the interplay between excitation and inhibition, and abnormal neurotransmitter transport has been shown to cause nerve dysfunction and damage (Liu and Edwards, 1997; Morrow *et al.*, 1998). *Glycine transporter type 2* (*GlyT2*) is a member of the Na^+/Cl^- dependent family of neurotransmitter transporters and is one of two glycine transporters identified in humans. It is involved in the reuptake of glycine at the synaptic cleft through a ligand-gated Cl^- channel (Amara & Kuhar, 1993 cited in Aubrey *et al.*, 2000; Morrow *et al.*, 1998).

Glycine and γ-aminobutyric acid (GABA), the major inhibitory neurotransmitters of the vertebrate brain stem and spinal cord, keep the postsynaptic membrane polarized, thereby suppressing synapse firing. In contrast, glutamate and other excitatory neurotransmitters function to open Na^+ channels and depolarize the presynaptic membrane until the action potential is reached and the synapse fires

(Reckling *et al.*, 2000; Alberts *et al.*, 1996). GABA and glycine are both considered fast neurotransmitters because of their kinetics; however, the kinetics of GABA suppression is different from glycine suppression, since GABA has a more prolonged effect than glycine (Nicoll and Malenka, 1998). Although they have separate receptors and transporters, GABA and glycine are co-released from the same vesicles at the inhibitory synapses of spinal neurons (Jonas *et al.*, 1998).

The GLYT2 protein contains six consensus protein kinase C-dependent phosphorylation sites as well as five *N*-linked glycosylation sites (Martinez-Maza *et al.*, 2001; Morrow *et al.*, 1998). It has been demonstrated that the glycosylation status of the protein affects its subcellular localization (Martinez-Maza *et al.*, 2001). Like other members of the Na^+/Cl^- dependent neurotransmitter transporter family, GLYT2 contains 12 transmembrane domains, with domain III being primarily responsible for the transport activity of both the substrate and the co-transported Na^+ and Cl^- ions (Ponce *et al.*, 2000). GLYT2 co-localizes with the strychnine-sensitive glycine receptors in the brain stem, spinal cord, and cerebellum (Morrow *et al.*, 1998; Aubrey *et al.*, 2000). It co-immunoprecipates with Syntaxin 1A, a SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor) protein involved in both intracellular membrane trafficking and the regulation of various ion channels. Syntaxin 1A has also been shown to interact with GAT1, a GABA transporter, and GLYT1, a type-1 glycine transporter. When GLYT2 is bound to syntaxin 1A, the normal GLYT2 protein trafficking pattern in the cell is upset. This causes GlyT2 to be improperly distributed in the plasma membrane, thereby interfering with the transport of glycine (Geerlings *et*

al., 2000). Furthermore, a yeast two-hybrid analysis using GlyT2 as bait has identifed unc-33 like phosphoprotein 6 (Ulip6) as a protein interacting with GLYT2. Ulip6 is expressed in embryonic and postnatal brain and spinal cord and is believed to be involved in neuronal differentiation and axon growth (Horiuchi *et. al.*, 2000). Therefore, GLYT2 is a protein involved in neural development and function in several ways.

In humans, the *GLYT2* gene produces a single, 9.5-kb mRNA transcript detected by northern hybridization in medulla, cerebellum, and spinal cord (Morrow *et al.*, 1998). The gene maps to human chromosome 11p15.1-15.2 (Morrow *et al.*, 1998), which is homologous to the *psrt* region of Mmu 7. Several polymorphisms have been reported in the human gene; however, the effect(s) of these polymorphisms on neurophysiology is unknown (Evans *et al.*, 1999).

Protein arginine methyltransferase 3 (*Prmt3*)

A thorough discussion of *Prmt3* as a candidate for the *psrt* phenotype should include an overview of the process of methylation. What follows is such an overview, with a description of enzymatic methylation in general followed by a review of protein arginine methylation, including a discussion of the functional roles of the process at the cellular and organismal levels, as well as *in vivo* substrates and protein arginine methyltransferase enzymes.

1. Enzymatic Methylation

Methyltransferases are an important class of enzymes that catalyze the transfer of a methyl group to a substrate. The universal methyl donor for these reactions is *S*-adenosylmethionine (AdoMet, or SAM), which is a methionine with an adenosine attached to the sulfur atom on the side chain; however, there are many types of methyl-accepting substrates. DNA methyltransferases are a well-studied class of enzymes responsible for methylating various positions on cytosine and adenine nucleotides (Bussiere *et al.*, 1998; Labahn *et al.*, 1994). A methyltransferase that modifies mRNA, and two small molecule methyltransferases that modify catechol and glycine, have also been described (Bussiere *et al.*, 1998). Several types of protein methyltransferases are known, as well (Gary and Clarke, 1998).

Despite the diversity of functions performed and substrates utilized by the different methyltransferase enzymes, all known AdoMet-dependent methyltransferases operate via a S_N2 (bimolecular nucleophilic substitution) reaction mechanism (Zhang *et al.*, 2000). They contain a signature methyltransferase domain composed of four conserved regions of homology: motifs I, post-I, II, and III. These motifs are responsible for binding AdoMet and catalyzing the methyl transfer (Weiss *et al.*, 2000). They appear in the same order and are separated by a characteristic number of residues. Motifs I and post-I are involved in AdoMet binding, motifs II and III interact with each other, and motifs I, II, and III are involved in forming the central β-sheet, a structural element common to all methyltransferases (Niewmierzycka and Clarke, 1999).

2. Protein Methylation

In eukaryotes, proteins can be methylated either on the oxygen molecule in the carboxyl group or on the nitrogen molecule(s) in the side chain. Only three amino acid residues are modified in eukaryotes by methylation: arginine, histidine, or lysine (Aletta *et al.*, 1998; Sundarraj and Pfeiffer, 1973; Gary and Clarke, 1998). In prokaryotes, it has been shown that a methyl group can be transferred via an esterification reaction to the γ-carboxyl group of glutamic acid residues (Aletta *et al.*, 1998).

There have been three types of protein methylases described: those that transfer the methyl group (from AdoMet) to an arginine residue in the substrate, those that transfer it to a C-terminal carboxyl group of a histidine residue, and those that transfer the methyl group to the ε-amino group of a lysine residue. These are called protein methylase I, II, and III, respectively (Sundarraj and Pfeiffer, 1973) (Figure VI-1).

Protein methylase I, a protein arginine methyltransferase (PRMT), is the enzyme responsible for methylating any of the guanidino nitrogens present in arginine residues. PRMT genes have been identified and cloned in several organisms, including human, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, *Arabidopsis*, and *Caenorhabditis elegans*. In addition, expressed sequence tags (ESTs) from at least six other organisms (including *Drosophila melanogaster*, *Xenopus laevis*, sea urchin and zebrafish, as well as rice and tomato) show similarity to PRMTs (Zhang *et al.*, 2000).

3. Functional Roles of Arginine Methylation

The precise biochemical role that protein arginine methylation plays in the cell has not yet been determined; however, it is believed to be responsible, at least in part, for a variety of processes. Most of the known *in vivo* targets of protein arginine methylation are proteins involved in some way with RNA, either through directly binding the RNA or through involvement with the ribosome (Najbauer *et al.*, 1993; Liu and Dreyfuss, 1995; Gary and Clarke, 1998). Certain proteins contain an RGG (or GAR) domain, which has an abundance of arginine and glycine residues and is the specific substrate for several protein arginine *N*-methyltransferases. This domain has been shown to modulate protein-protein and protein-nucleic acid interactions (Bouvet *et al.*, 1998; Gary and Clarke, 1998). Methylarginine-containing domains have also been shown to be responsible for the binding specificity of certain proteins (Plotnikov *et al.*, 2000). It has been postulated that the presence of methylarginine residues could affect the dynamics of ligand-substrate interactions, altering either the stereochemistry of the process, the hydrogen bonding interactions occurring between the modified protein and its substrate(s), or both (Mears and Rice, 1996). Therefore, the methylation of arginine residues is believed to play a role in the intermolecular dynamics of ligand-substrate binding, including both protein-nucleic acid and protein-protein interactions.

The modification of arginine residues has also been shown to affect a protein's susceptiblilty to proteolysis. Methylated heterogeneous nuclear ribonuclear protein A1 (hnRNP A1) is more susceptible to trypsin digestion than the unmethylated protein (Rajpurohit *et al.*, 1994a), while the methylated forms of other hnRNPs have been

shown to confer resistance to proteolytic cleavage (Bedford *et al.*, 2000). The myelin basic protein (MBP) normally contains 19 arginine residues, only one of which is modified by methylation. It has been shown that MBP containing this single methylarginine residue is more resistant to trypsin digestion than is the unmethylated MBP (Brostoff and Eylar, 1971). On the other hand, MBP containing citrullines in place of the methylarginine and the other unmodified arginines is proteolytically degraded by cathepsin D more quickly than the non-citrullated MBP. Citrulline, like methylarginine, is a posttranslationally modified arginine; however, citrulline is modified by deamidation, a different process. Furthermore, there is a positive correlation between number of citrulline residues in the protein and the rate of enzymatic degradation (Cao *et al.*, 1999). Therefore, the modification of arginine residues, either through the process of protein arginine methylation or through deamidation, can effectively and substantially alter the protein-protein interactions involved in catalytic proteolysis.

Protein arginine methylation has also been shown to be involved in subcellular localization. Proteins containing methylarginine residues are found in various compartments within the cell: some are located in the nucleus, some are cytosolic, and others are capable of shuttling between the nucleus or nucleolus and the cytoplasm. It has been demonstrated that the methylarginine-containing domain of certain proteins is necessary for them to be localized to the proper subcellular compartment (Schmidt-Zachmann and Nigg, 1993; Burgess *et al.*, 1991; Gary and Clarke, 1998; Nichols *et al.*, 2000). Similarly, it has been shown in *Saccharomyces cerevisiae* that protein arginine

methylation is required for certain hnRNPs to be exported from the nucleus to the cytoplasm (Shen *et al.*, 1998).

It has long been reported that the posttranslational acetylation of certain lysine residues in the histones is responsible for regulating gene expression. The acetylation of the histones effects a change in the structure of chromatin, causing it to "loosen", thereby allowing the cellular machinery easier access to DNA (Pennisi, 1997; Rundlett *et al.*, 1996). It has been suggested that a similar posttranslational modification, the methylation of arginine residues in the histone proteins, may also be at least partially responsible for altering the structure of chromatin (Chen *et al.*, 1999; Hagmann, 1999). There have been numerous protein deacetylases identified and studied in eukaryotes; however, no protein arginine demethylase has yet been identified (Smith *et al.*, 1999), although the existence of one has been proposed (Young and Waickus, 1987). Thus it appears that methylation, unlike acetylation, may be an irreversible process. Therefore, gene regulation through histone methylation may or may not follow the same paradigm as gene regulation through reversible acetylation. It should be noted, however, that the histones are complex proteins that are the targets of numerous epigenetic modifications (Hagmann, 1999); therefore, it will be difficult to determine the exact effect(s) that protein arginine methylation may have on these proteins, especially since these effects could be subtle.

Along these same lines, there is some evidence to indicate that protein arginine methylation may be connected to development. In the rat, methyltransferase activity appears to be developmentally regulated. Protein arginine methyltransferase activity is

highest in fetal brain, declining steadily after birth (Paik *et al.*, 1972 cited in Gary and Clarke, 1998). These results suggest a possible connection between protein arginine methylation and development, either in the developmentally regulated expression of genes or in some other aspect of growth and development.

It has also been hypothesized that there may be a link between protein arginine methylation and metastasis (Gu *et al.*, 1999; Pennisi, 1997). Elevated levels of protein arginine *N*-methyltransferase activity have been reported in highly proliferating tissues, and there is a correlation between the level of methyltransferase activity and the degree of proliferation (Rajpurohit *et al.*, 1994b; Scorilas *et al.*, 2000). While there are three mRNA transcripts of protein arginine methyltransferase 1 (*Prmt1*) present in normal mammary cells, two of these transcripts are significantly down-regulated in breast cancer cells (Scorilas *et al.*, 2000). In addition, a 20 kDa protein of unknown function has been isolated from four different cancer cell lines and shown to be methylated at multiple arginine residues; however, this same protein is present but unmethylated in normal colon cells (Gu *et al.*, 1999). Although not conclusive, this evidence suggests the involvement of protein arginine methylation in cellular proliferation and metastasis.

The human Janus kinase binding protein 1 (JBP1) and its yeast homolog, Hsl7p, contain a methyltransferase motif and have been shown to exhibit protein arginine methyltransferase activity (Pollack *et al.*, 1999; Lee *et al.*, 2000). JBP1 binds to Janus kinase (Jak), which is involved in the Jak-Stat (signal transducers and activators of transcription) pathway (Pollack *et al.*, 1999). PRMT1, a mammalian protein arginine methyltransferase, was first identified because of its ability to interact with the

intracytoplasmic domain of the interferon-α,β receptor, which is also involved in the Jak-Stat pathway (Abramovich *et al.*, 1997). The Jak-Stat pathway is involved in both signal transduction and cell cycle regulation, suggesting that methyltransferases may play a role in cell signaling and/or cell cycle regulation (Pollack *et al.*, 1999).

The methylation of a single arginine residue, arginine-107 (R-107), in the myelin basic protein (MBP) is critical to the function of MBP within the myelin membrane. It has been shown that arginine methylation is involved in maintaining the integrity of the myelin sheath, and the inhibition of MBP-specific methyltransferase activity during myelinogenesis leads to dysmyelination (Amur *et al.*, 1986; Kim *et al.*, 1997b). MBP undergoes several posttranslational modifications, all of which affect the charge microheterogeneity of the mature protein. These modifications, one of which is methylation, alter the hydrophobicity of the protein which, in turn, affects the ability of the MBP to interact with, and bind to, the lipids in the myelin membrane (Zand *et al.*, 1998; Boggs *et al.*, 1999). Methylated MBP is less cationic than unmethylated MBP (Ghosh *et al.*, 1991), and the methylation of MBP alters the thermodynamics of the MBP-myelin membrane association by decreasing the free energy of the protein-lipid interaction by 0.9 kcal/mole (Young *et al.*, 1987). The methylarginine residue in MBP occurs near the N-terminus, and it has been shown *in vitro* that the N-terminal half of the MBP interacts with lipids more than the C-terminal half of the protein (Boggs *et al.*, 1999). Therefore, protein arginine methylation is involved in the development and maintenance of the mammalian central nervous system.

Protein arginine methylation has also been implicated in a number of autoimmune diseases. There has been some evidence presented to link the methylation of R-107 in the MBP to multiple sclerosis (MS). MS is a neurological disease characterized by an erosion of the myelin sheath and is believed to be caused by an autoimmune attack on myelin. It has been suggested that the autoantigen responsible for the disease is MBP itself (Moscarello *et al.*, 1994). Once the myelin has been disrupted, MBP is broken down in the body and excreted in the urine (Whitaker, 1998; Whitaker *et al.*, 1999). Analysis of the dimethylarginine (Me_2Arg) content of MBP breakdown products in the urine of chronic MS sufferers confirms the presence of 20-33% less Me_2(sym)Arg than would be expected (Rawal et al., 1995b). Since methylated MBP is resistant to enzymatic digestion, it has been suggested that the rate of MBP degradation and, therefore, the progression of the disease, may be related to the methylation status of the R-107 in the protein (Kim *et al.*, 1997b).

It has also been shown that, based on the morphology of the myelin membrane and the level of charge microheterogeneity in the MBP, the myelin present in MS patients is developmentally immature (Moscarello *et al.*, 1994). The composition of the MBP in myelin changes throughout development. In the mouse and other rodents, where several alternatively spliced isoforms of MBP exist, the temporal expression pattern of each isoform is unique (Barbarese *et al.*, 1978; Carson *et al.*, 1983; Kruger *et al.*, 1999). It has also been shown that the amount and type of MBP methylation in young rodent brains is different from that found in mature brains (Rawal et al., 1995a). MBP isolated from normal adult rat brains contains MeArg and symmetrical

dimethylarginine (Me_2(sym)Arg). In young rat brains, however, MBP contains an additional isomer, asymmetrical dimethylarginine (Me_2(asym)Arg), which is not found in adult brains (Rawal et al., 1995a). It is not clear if the immature myelin present in MS patients is either the cause, or the result, of the disease; in either case, the etiology of the disease includes demyelination, which may be linked to abnormal MBP methylation.

In addition to its possible role in MS, MBP is the primary autoantigen associated with allergic encephalomyelitis (Eylar and Thompson, 1969). There are other methylarginine containing proteins that have been associated with autoimmune diseases, as well. Fibrillarin, which contains several methylarginine residues, was first identified because of its immunoreactivity to scleroderma antibodies (Ochs *et al.*, 1985). Antibodies from patients with systemic lupus erythematosus have been used to identify three methylarginine-containing protein autoantigens involved in this disease: heterogeneous ribonucleoprotein A1, nucleolin, and Sm protein D1 (Minota *et al.*, 1991; Jensen *et al.*, 1988; Brahms *et al.*, 2000). On a separate, but related, note: peptides containing citrulline, a different posttranslational arginine modification, have been implicated in rheumatoid arthritis, another autoimmune disease (Schellekens *et al.*, 1998). Therefore, there appears to be a link between arginine modification and the etiology of certain autoimmune diseases.

Protein arginine methylation is an important posttranslational modification that is believed to play a variety of roles in the cell and in the mammalian central nervous and immune systems. Proteins containing methylarginine residues are associated with

substrate-ligand binding, including protein-nucleic acid and protein–protein interactions. They are also associated with signal transduction, development, the regulation of gene expression and cellular proliferation, nervous system function, and the etiology of certain autoimmune diseases. In addition, methylarginine-containing domains have been shown to be necessary for the proper subcellular localization of many proteins. Although the precise role(s) that protein arginine methylation plays has yet to be definitively determined, it is clear that the processes associated with this modification are numerous and diverse, and necessary for proper cellular and organismal function.

4. *In vivo* Substrates for Protein Arginine Methyltransferases

There are several known *in vivo* substrates for protein arginine methyltransferases. Some are localized in the nucleus or nucleolus, some are cytosolic, and others are capable of shuttling between the nucleus/nucleolus and the cytoplasm. The biological function of many of these substrates is well known, while the role of others has yet to be elucidated. Most of the known substrates are either monomethylated or asymmetrically dimethylated; however, two substrates have been identified that contain symmetrically dimethylated arginine residues.

A. Asymmetrically Methylated Substrates

1. Yeast Hrp1p Protein

Heterogeneous nuclear ribonuclear proteins (hnRNPs) are RNA-binding proteins involved in the complex process of RNA maturation. There are at least 20

hnRNPs in the eukaryotic nucleus, and most of these have been shown to contain methylarginine residues (Liu and Dreyfuss, 1995). The *Saccharomyces cerevisiae* Hrp1p protein is an hnRNP involved in the polyadenylation and processing of the 3' end of mRNAs as well as mRNA export from the nucleus. Although it is predominately located in the nucleus, it has been found to shuttle between the nucleus and the cytoplasm. Hrp1p contains asymmetrically dimethylated arginine residues; however, it has been shown that the methylation of the arginine residues has little or no effect on its ability to bind RNA (Valentini *et al.*, 1999).

2. Heterogeneous Nuclear Ribonuclear Protein A1

The heterogeneous nuclear ribonuclear protein A1 (hnRNP A1) is one of the most abundant of the multiple polypeptide components of the eukaryotic 40S ribonucleoprotein particle. It contains a unique nuclear localization signal sequence that also functions as a nuclear export signal. Although hnRNP A1 is primarily localized in the nucleus, it shuttles between the nucleus and the cytoplasm via a temperature-sensitive pathway. Therefore, it is believed that hnRNP A1 is involved in shuttling mRNA molecules from the nucleus to the cytoplasm (Michael *et al.*, 1995). Alternatively spliced transcripts for the human HNRNPA1 gene can encode for at least two protein isoforms (Buvoli *et al.*, 1988; Biamonti *et al.*, 1989). The protein itself contains a glycine- and arginine-rich (GAR) region, with four known sites for protein arginine methylation near the C-terminus (Williams *et al.*, 1985; Rajpurohit *et al.*, 1994b; Kim *et al.*, 1997c). An analysis of the modified arginine residues from this

region reveal both monomethylarginine and asymmetrically dimethylated arginine (Rajpurohit *et al.*, 1994b).

3. Poly(A)-binding Protein II

Poly(A)-binding protein II (PABP2) is a protein that binds to the polyadenylated tails of nascent mRNA transcripts. Although found predominately in the nucleus, it has the capability to shuttle between the nucleus and the cytoplasm (Calado *et al.*, 2000; Smith *et al.*, 1999). It contains an RNA-binding domain, an acidic N-terminus, and a basic C-terminus containing 13 asymmetrically dimethylated and two monomethylated arginine residues (Nemeth *et al.*, 1995; Smith *et al.*, 1999). In the rat, the *Pabp2* gene produces a 9.5-kb transcript that is expressed in brain, kidney, liver, lung, and testis (Muller *et al.*, 1992). Mutations involving expansion repeats in the human *PABP2* gene cause the autosomal dominant form of oculopharyngeal muscular dystrophy (OPMD), a late-onset neurological disease (Calado *et al.*, 2000).

4. Fibrillarin

Fibrillarin is a nucleolar protein that is a part of the small nuclear ribonucleoprotein particle (snRNP). It is believed to be involved in processing RNA, specifically by shortening the preribosomal RNA from the 5' end (Aris and Blobel, 1991). Fibrillarin contains at least six arginine residues that have been shown to be asymmetrically dimethylated, all of which are contained in the GAR region of the N-terminal domain (Gary and Clarke, 1998; Aris and Blobel, 1991). Fibrillarin is present

in all eukaryotic species examined thus far, and fibrillarin genes have been cloned in *Saccharomyces cerevisiae* and *Xenopus laevis* (Aris and Blobel, 1991).

5. Nucleolin and U20

Nucleolin, like fibrillarin, is a localized in the nucleolus. It is a phosphoprotein involved in the biogenesis and maturation of ribosomes (Srivastava *et al.*, 1990). Unlike fibrillarin, however, nucleolin has the ability to shuttle between the cytoplasm and the nucleus. For this reason, it is believed that nucleolin is involved not only in the assembly of the ribosomes, but also in the export of pre-ribosomal particles from the nucleolus to the cytoplasm. The C-terminus of nucleolin contains an RGG domain, which has an abundance of arginine and glycine residues. This RGG domain, which contains asymmetrically dimethylated arginine residues, functions to modulate interactions between nucleolin and several ribosomal proteins (Bouvet *et al.*, 1998). It is also required for proper subcellular localization (Schmidt-Zachmann and Nigg, 1993). The nucleolin gene is well conserved across several species; it is present in chicken and yeast as well as in several mammalian species, where it is ubiquitously and constitutively expressed (Srivastava *et al.*, 1990). The nucleolin gene is also responsible for producing U20, a short (80 nt) RNA transcript which, like nucleolin, is localized in the nucleus. U20 is immunoprecipated by antibodies to fibrillarin; therefore, fibrillarin and U20 may interact *in vivo*. U20 transcripts, which contain a 21 nt stretch that is complementary to the sequence of the 18S rRNA, have been detected in all vertebrates examined so far (Nicoloso *et al.*, 1994).

6. Histones

The histone proteins are another class of proteins that have been shown to be *in vivo* substrates for protein methyltransferases (Casellas and Jeanteur, 1978). Specific amino acid residues on the histones can be modified by various biological processes, including phosphorylation, acetylation and methylation. Although the histones are normally methylated on lysine residues by a protein lysine methyltransferase, it has been shown that arginine residues on various histone proteins are substrates for protein arginine methyltransferases (Hagmann, 1999). Arginine residues on histones H1, H2, H3, and H4 can be methylated (Frost *et al.*, 1989; Gary and Clarke, 1998; Chen *et al.*, 1999; Disa *et al.*, 1986) even though histones H2B, H2A and H3 are devoid of the GAR motif that has been shown to be the target of arginine methylation in other proteins. These non-GAR histones do, however, contain multiple arginine residues that could serve as methyl acceptors (Gary and Clarke, 1998). There have been conflicting reports about the product of protein arginine methylation in the histones; therefore, it is unclear if the *in vivo* product of this process is monomethylarginine, asymmetrically dimethylated arginine, symmetrically dimethylated arginine, or some combination of the three species (Gary and Clarke, 1998).

7. Basic Fibroblast Growth Factor

Basic fibroblast growth factor (bFGF), or FGF-2, is a mitogen present in mammalian cells (Gary and Clarke, 1998). FGF-2 is a heparin-binding protein that is believed to be involved in angiogenesis, osteogenesis, tissue repair, and blood pressure

regulation (Abraham *et al.*, 1986; Burgess *et al.*, 1991; Dono *et al.*, 1998; Montero *et al.*, 2000). Mice lacking Fgf-2 display a complex phenotype indicative of the pleiotropic nature of the protein. Although Fgf-2 null mice are viable, fertile, and appear morphologically normal, they display structural defects in several areas of the central nervous system, including the cerebral cortex, the hippocampus, and the cervical spinal cord. In addition, Fgf-2 deficient mice are hypotensive and have a retarded healing response to skin injury (Dono *et al.*, 1998; Ortega *et al.*, 1998). The *Fgf-2* gene encodes multiple transcripts through alternative splicing events (Kurokawa *et al.*, 1987). The high molecular weight (HMW) isoforms of the FGF-2 protein have an amino terminal extension containing an RGG domain with multiple sites for arginine methylation, and the arginine residues in this region are asymmetrically dimethylated (Klein *et al.*, 2000; Gary and Clarke, 1998; Burgess *et al.*, 1991). It has been suggested that this N-terminal domain, containing the methylarginine residues, is responsible for both the nuclear localization and the binding specificity of the HMW FGF-2 isoforms (Burgess *et al.*, 1991; Plotnikov *et al.*, 2000).

8. ICP27

The ICP27 protein encoded by the herpes simplex virus type 1 (HSV-1) contains one or more methylarginine residues in a glycine- and arginine-rich motif similar to an RGG box. ICP27 is an immediate-early (IE, or α) protein that is primarily localized in the nucleus; however, it has been shown to shuttle between the nucleolus and the cytoplasm (Mears and Rice, 1996; Sandri-Goldin, 1998). ICP27 is involved in the

polyadenylation and splicing of RNA; it stimulates 3' RNA processing of polyadenylated viral proteins while inhibiting the splicing of host mRNAs (Sandri-Goldin, 1994). The RGG-like motif in ICP27, where the methylarginine(s) is/are located, is required for the protein to bind to RNA (Sandri-Goldin, 1998). Although this motif is not necessary for nuclear localization, it is required for the proper localization to the nucleolus (Mears *et al.*, 1995).

9. Other Proteins

There are several other proteins that function as substrates for Type I protein arginine methyltransferases *in vivo*. Interleukin enhancer-binding factor 3 (ILF3) is a protein whose C-terminal is rich in glycine, arginine, and serine residues. It is a nuclear protein that is a substrate for, and co-immunoprecipitates with, PRMT1 (Tang *et al.*, 2000). The yeast essential Npl3p (nucleosome assembly protein 1-like 3) protein is an RNA binding protein that shuttles between the nucleus and the cytoplasm. It contains multiple RGG motifs in its C-terminus, and evidence suggests that these C-terminal RGG motifs contains one or more methylarginine residues (Henry and Silver, 1996; Siebel and Guthrie, 1996). Sam68, a proposed adapter protein for Src kinases, contains a proline-rich motif flanked by a GAR domain containing asymmetrically dimethylated arginine residues (Bedford *et al.*, 2000). Cytochrome c from *Euglena gracilis* is a cytosolic protein containing a single methylarginine residue near the N-terminus (Farooqui *et al.*, 1985). In addition, Garl (glutamic acid-rich protein 1), Nopl (nucleolar protein 1), and Sbpl (skeleton binding protein 1) are all GAR-containing

proteins which have been shown to be methyl acceptors *in vivo* (Frankel and Clarke, 1999).

B. Symmetrically Methylated Substrates

1. Sm Proteins D1 and D3

The Sm proteins D1 and D3 are two of the numerous protein constituents of the spliceosome, a nuclear protein complex involved in the processing and splicing of pre-mRNA molecules. In all eukaryotes except yeast, these proteins contain RG dipeptide repeats in their C-termini. These RG dipeptides, which are different from the RGG or GAR motifs found in other methylated substrates of PRMTs, are the site of *in vivo* arginine methylation. The human D1 and D3 proteins contain nine and four arginine residues, respectively, in their C-terminal RG dipeptide motif, and all of these arginine residues exist exclusively as symmetrical dimethylarginine (Brahms *et al.*, 2000).

2. Myelin Basic Protein

The myelin basic protein (MBP) is one of the major protein constituents of the myelin membrane. It is a highly basic protein and is the antigen responsible for allergic encephalomyelitis (Eylar and Thompson, 1969). In solution, the MBP molecule has very little secondary structure. It is considered to be highly unfolded with an open and extended rod-like conformation, and it is this lack of secondary structure that confers its unusual resistance to denaturation (Eylar and Thompson, 1969). MBP interacts *in vivo*

with acidic lipids to form the myelin sheath, and this lipid-bound state induces additional structural characteristics in the molecule (Lees and Brostoff, 1984).

Although MBP exists as a single protein in humans (Carnegie, 1970 cited in Kim *et al.*, 1997b), the MBP of rodents is actually several related proteins arising from alternatively spliced transcripts of a single gene (deFerra *et al.*, 1985; Takahashi *et al.*, 1985). There are four isoforms of MBP found in the central nervous system (CNS) tissue of mice, ranging in mass from 14 kDa to 21.5 kDa (deFerra *et al.*, 1985), with each isoform having a unique expression pattern (Barbarese *et al.*, 1978; Carson *et al.*, 1983; Kruger *et al.*, 1999). In the mouse, MBP mRNA expression, MBP protein synthesis, and myelination occur synchronously (Campagnoni *et al.*, 1978; Carson *et al.*, 1983).

The MBP gene is a part of the larger Golli (gene expressed in the oligodendrocyte lineage) transcription unit. The human and mouse Golli-MBP genes have been cloned and sequenced; they contain 10 exons, the last 7 of which encode for the various isoforms of MBP (Pribyl *et al.*, 1993; Campagnoni *et al.*, 1993). The Golli-MBP gene maps to human chromosome 18q22-q23 (Kamholz *et al*, 1987); coincidentally, the mouse homolog maps to Mmu 18 (Sidman, 1985). The MBP gene contains a regulatory element located 50 nt upstream of the start of MBP transcription (in exon 3 of the Golli-MBP gene), and this regulatory element is required for proper temporal expression of the gene (Haas *et al.*, 1995). There is also a RTS (mRNA transport sequence) in the 3' untranslated region where hnRNPs bind and transport the MBP mRNA to the distal end of the oligodendrocyte processes. Once there, it is

incorporated into the myelin membrane and translated into protein (Hoek *et al.*, 1998). In the absence of thyroid hormone (which is involved in the regulation of myelinogenesis), 9-*cis*-retinoic acid, the ligand for the retinoic acid receptor, can stimulate transcription of the MBP promoter in rodent neural cell cultures (Pombo *et al.*, 1999).

MBP may serve a variety of functions *in vivo*. In addition to its well-characterized role in myelinogenesis, MBP can catalyze the polymerization and assembly of clathrin baskets *in vitro* (Prasad *et al.*, 1995). Human, bovine and rat MBPs have all been shown to stimulate insulin and glucagon secretion from rat pancreatic islets (Kolehmainen *et al.*, 1990), and bovine MBP can inhibit the activity of the histone-specific PRMT (Park *et al.*, 1989). It has also been shown that bovine MBP, when injected in Djungarian hamsters, causes a variety of effects in the endocrine pancreas, associating with, and disrupting the integrity of, the plasma and intracellular membranes (Kolehmainen and Sormunen, 1998).

MBP is a highly modified protein *in vivo*, and these post-translational modifications serve to introduce charge microheterogeneity in the protein. Myelin isolated from CNS tissue contains any number of the following modifications: phosphorylated serine and threonine residues; an acetylated *N*-terminal residue; the deamidation of either a specific glutamine residue or any of the 19 arginine residues; and the methylation of a single arginine residue to produce both monomethylarginine and symmetrically dimethylated arginine (Zand *et al.*, 1998; Boggs *et al.*, 1999; Young *et al.*, 1987; Baldwin and Carnegie, 1971a; Deibler and Martinson, 1973; Gary and

Clarke, 1998). The methylation of the arginine-107 residue appears to be a nearly universal modification: with the exception of carp, it has been found in every mammalian and sub-mammalian vertebrate species studied (Kies *et al.*, 1972; Deibler and Martenson, 1973; Young *et al.*, 1987).

There are several known *in vivo* substrates for protein arginine methyltransferases. These posttranslationally modified proteins are involved in a number of cellular processes, playing a variety of diverse roles. However, since the *in vivo* substrates of several PRMTs are unknown, there are likely to be additional substrates added to this list. Identification and comparison of these substrates will undoubtedly lead to a deeper understanding of the mechanism of the PRMT reaction and the role(s) that protein arginine methylation plays in the cell.

4. Protein *N*-Arginine Methyltransferases (PRMTs)

PRMT enzymes can be categorized into three subtypes: those that produce ω-N^G-monomethylarginine and (asymmetric) ω-N^G,N^G-dimethylarginine (often referred to as Type I); those that produce ω-N^G-monomethylarginine and (symmetric) ω-N^G,N'^G-dimethylarginine (Type II); and those that produce δ-N^G-monomethylarginine (Gary and Clarke, 1998; Zobel-Thropp *et al.*, 1998; Niewmierzycka and Clarke, 1999). The third type of PRMT has only recently been identified and the protein is not well characterized (Niewmierzycka and Clarke, 1999; Zobel-Thropp *et al.*, 1998) (Figure VI-1).

Both Type I and Type II PRMTs are stimulated by cyclic AMP and inhibited by S-adenosylhomocysteine (one of the products of the methylation reaction), sinefungin (or adenosyl ornithine, an antibiotic), adenosine dialdehyde, and catecholamines such as epinephrine, norepinephrine, isoproterenol (Amur *et al.*, 1986; Najbauer *et al.*, 1992; Park *et al.*, 1989; Bussiere *et al.*, 1998). The reactions involving these enzymes are highly endothermic, requiring 12 ATPs per methyl group attached, and are apparently irreversible (Gary and Clarke, 1998).

A. Type I PRMTs

There have been several Type I PRMTs identified and cloned in a wide variety of eukaryotic organisms. All of them contain the signature methyltransferase motifs; however, these enzymes can, and do, vary widely in their sequence outside these regions. There are only three PRMTs known so far in yeast, all of which have homologs in other eukaryotes. There are five PRMTs cloned so far in mammals, and it is possible that there are others that have not yet been identified.

1. Yeast Type I PRMTs

There have been two Type I methyltransferases identified from *Saccharomyces cerevisiae*: RMT1 (protein arginine methyltransferase 1) and RRP8. RMT1, which has also been called HMT1 (hnRNP methyltransferase 1), was identified by its homology to the rat PRMT1 gene (Gary *et al.*, 1996; Gary and Clarke, 1998). RMT1 contains four conserved methyltransferase motifs: I, post I, II, and III (Gary *et al.*, 1996). It is a nuclear protein that can methylate a number of substrates *in vivo*, including NPL3

(nucleosome assembly protein 1-like 3), a yeast hnRNP. RMT1, unlike its NPL3 substrate, is not essential for cell viability (Henry and Silver, 1996). Crystallographic evidence suggests that RMT1 is oligomeric, existing as a hexamer in solution (Weiss *et al.*, 2000). Sequence analysis of the *Rmt1* gene reveals two regions of similarity to other protein arginine methyltransferases: a stretch of 63 nucleotides showing strong similarity to the SAM-binding motif of other methyltransferases, and a stretch of 48 nucleotides containing a nuclear localization signal (Henry and Silver, 1996).

RRP8 is another methyltransferase identified from *Saccharomyces cerevisiae*. It is a nucleolar protein involved in pre-rRNA cleavage. Database searches indicate that the C-terminal domain of RRP8 shows homology to proteins from a variety of organisms, including *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Plasmodium falciparum*, *Schizosaccharomyces pombe*, and mammalian species as well (Bousquet-Antonelli *et al.*, 2000). RRP8 was first reported to contain only one signature methyltransferase domain (Niewmierzycka and Clarke, 1999); however, further analysis confirms that it contains four conserved methyltransferase domains: I, post I, II, and III (Bousquet-Antonelli *et al.*, 2000).

2. Mammalian Type I PRMTs

a. Protein arginine methyltransferase 1

Protein arginine methyltransferase 1 (PRMT1) is the first mammalian PRMT to be identified and cloned, and is responsible for the majority of PRMT enzyme activity found in mammals (Lin *et al.*, 1996; Tang *et al.*, 2000). It was recovered using a yeast

two-hybrid assay to identify proteins interacting with two similar proteins: TIS21, an immediate-early protein of unknown function, and BTG1, an antiproliferative protein involved in chronic lymphocytic leukemia (Lin *et al.*, 1996). PRMT1 was later identified as a ligand for the Interleukin enhancer-binding factor 3 (ILF3), and has been shown to methylate the C-terminus of ILF3 *in vitro* (Tang *et al.*, 2000). It has also been shown to interact with both the intracytoplasmic domain of the interferon-α,β receptor and the C-terminal AD2 activation domain of p160 family of coactivators (Abramovich *et al.*, 1997; Koh *et al.*, 2001).

PRMT1 is classified as a type I protein arginine *N*-methyltransferase because of its ability to form monomethyl- and asymmetrical dimethylarginine residues (Gary and Clarke, 1998). The *in vitro* activity of PRMT1 is strongly inhibited in the presence of Tris and arginine but is unaffected by citrulline (a deamidated arginine), or the steroid hormones (Frankel and Clarke, 2000; Scorilas *et al.*, 2000). Conversely, ILF3 has been shown to enhance the activity of PRMT1 *in vitro* (Tang *et al.*, 2000). Native PRMT1 can methylate the histones, hnRNP A1, and the poly(A)-binding protein II (PABP2) (Lin *et al.*, 1996; Smith *et al.*, 1999). A Glutathione *S*-transferase (GST) fusion protein containing the catalytic portion of PRMT1 has been shown to methylate GST-GAR, an artificial substrate containing GST fused to the glycine- and arginine-rich amino-terminal residues of fibrillarin (Tang *et al.*, 1998). However, the GST-PRMT1 protein is unable to methylate either cytochrome *c* or the myelin basic protein (Lin *et al.*, 1996; Smith *et al.*, 1999).

PRMT1 contains four highly conserved methyltransferase motifs: I, post I, II, and III (Lin *et al.*, 1996). It is interesting to note, however, that PRMT1 contains few residues other than those identified in the conserved motifs (Frankel and Clarke, 1999). PRMT1 is an oligomeric protein, with reports of dimers and hexamers forming in solution (Tang *et al.*, 1998; Weiss *et al.*, 2000). Although the exact function of the PRMT1 oligomers is unknown, it has been shown that dimerization is essential for catalytic activity (Weiss *et al.*, 2000).

The PRMT1 gene contains 12 coding exons and spans an 11.2-kb genomic interval (Scorilas *et al.*, 2000). It maps to human chromosome 19q13.3 (Scott *et al.*, 1998; Scorilas *et al.*, 2000). In the mouse it maps to chromosome 7, proximal to *Ldh3*, *Gas2*, *Snrpn*, and *Ube3a* (Pawlak *et al.*, 2000). This places *Prmt1* just proximal to the *p*-deletion complex. There are no human disease phenotypes or mouse mutations known to map near these regions of the human or mouse genomes (Pawlak *et al.*, 2000).

Prmt1 produces a primary 1.282-kb mRNA transcript in the mouse (Pawlak *et al.*, 2000); however, at least two alternatively spliced transcripts, differing at their 5' ends, have been detected (Scott *et al.*, 1998; Scorilas *et al.*, 2000). *Prmt1* is ubiquitously expressed in both rodent and human tissues (Lin *et al.*, 1996; Scott *et al.*, 1998). Cell culture studies indicate that its activity is not required for cell viability; however, *Prmt1* is constitutively expressed in mitogen-stimulated cells and is necessary for early embryonic development in the mouse (Lin *et al.*, 1996; Pawlak *et al.*, 2000). While all three mRNA transcripts of Prmt1 are present in normal mammary cells, two

of the transcripts have been shown to be significantly down-regulated in breast cancer cells (Scorilas *et al.*, 2000).

Of all the known mammalian protein arginine methyltransferase genes, only one mutation, arising from a targeted retroviral insertion in the mouse *prmt1* gene, has been described. This *prmt1* mutation is essentially a null mutation, producing mRNA transcript levels of only 1%, compared to the wild-type. The *prmt1* knockout is embryonic lethal, with death occurring at around day E6.5. Although the *prmt1* null mutation is an embryonic lethal and PRMT1 is responsible for the major proportion of methyltransferase activity in cells, it has been shown that PRMT1 activity is not required for viability in mouse ES cells (Pawlak *et al.*, 2000).

b. Protein arginine methyltransferase 2

The human PRMT2 gene was identified because of its similarity to the rat PRMT1 gene. It maps to human chromosome 21q22.3 (Scott *et al.*, 1998; Katsanis *et al.*, 1997), and contains five conserved methyltransferase domains: I, post I, II, III, and post III. It also contains a Src homology 3 (SH3) domain in its 5' end, making it unique among all other previously identified members of the protein arginine methyltransferase family (Katsanis *et al.*, 1997). This SH3 domain is involved in protein-protein binding, and methylarginine residues present in the protein substrate serve to reduce the level of interaction between the SH3-containing ligand and its (methylarginine-containing) substrate (Bedford *et al.*, 2000). The Prmt2 gene encodes a 2.8-kb mRNA that is ubiquitously expressed, with slightly higher levels of expression in brain and placenta

and lower levels in the liver and lung. There is some evidence to suggest that a larger transcript may be present in breast and fetal heart, but this has not been conclusively determined (Katsanis *et al.*, 1997). It is interesting to note, however, that despite the homology between PRMT2 and other protein arginine methyltransferase family members, the PRMT2 protein has not been shown to have protein arginine methyltransferase activity (Scott *et al.*, 1998). Therefore, PRMT2 may be a structural, but not functional, member of the PRMT family.

c. Protein arginine methyltransferase 3

Protein arginine methyltransferase 3 (PRMT3) is a type I protein arginine *N*-methyltransferase that was recovered using a yeast two-hybrid assay to identify proteins interacting with PRMT1, the first mammalian PRMT to be identified and cloned. The PRMT1 and PRMT3 proteins are 67% similar and 46% identical at the amino acid level over the methyltransferase domain in their carboxy-terminal region; however, their amino-terminal ends are dissimilar. The N-terminus of PRMT3, which contains an abundance of acidic residues, is called the N-terminal acidic amino acid-rich (NAR) domain. This NAR domain, which contains both a C2H2 zinc finger motif and a tyrosine phosphorylation consensus sequence, shows no significant homology to other known proteins (Tang *et al.*, 1998).

The *in vivo* substrate for PRMT3 is unknown; however, a glutathione *S*-transferase (GST) fusion protein containing the catalytic portion of PRMT3 has been shown to methylate GST-GAR, an artificial substrate containing GST fused to the

glycine- and arginine-rich amino-terminal residues of fibrillarin, which has multiple arginine residues that serve as targets for methylation (Tang *et al.*, 1998). PRMT3 and PRMT1 have both been shown to methylate arginine residues in the C-terminal domain of poly(A)-binding protein II (PABP2), a protein believed to be involved in pre-mRNA polyadenylation, in a methylation pattern consistent with the natural methylation of PABP2 (Smith *et al.*, 1999). The activity of PRMT3, unlike that of PRMT1, is relatively unaffected by the addition of Tris to the *in vitro* reaction mixture; however, both enzymes are inhibited by arginine, and both are unaffected by the addition of citrulline, the posttranslationally deamidated adduct of arginine (Frankel and Clarke, 2000). In addition, there is a single 29 kDa substrate present in hypomethylated yeast *rmt1* extracts that is methylated by GST-PRMT3, but not GST-PRMT1 (Tang *et al.*, 1998).

PRMT3 is the only known enzyme with protein arginine *N*-methyltransferase activity to have a zinc finger motif. PRMT3 binds zinc in a 1:1 stoichiometric ratio, and it is this zinc finger domain of PRMT3 that appears to be responsible for conferring substrate specificity (Frankel and Clarke, 2000). GST-PRMT3 fusion proteins containing only the catalytic portion of PRMT3 (without the zinc finger domain) will methylate GST-GAR; however, it was found that the zinc-chelated form of PRMT3 was required for the enzyme to recognize certain RNA-associated substrates in RAT1 cell extracts (Frankel and Clarke, 2000). This is consistent with the notion that zinc fingers are involved in protein-ligand binding, including protein-DNA, protein-RNA, and protein-protein interactions (Mackay and Crossley, 1998).

The *Prmt3* gene produces a 2.4-kb mRNA transcript in the rat, which is constitutively and ubiquitously expressed except in liver, where no expression is detected via Northern blotting. Immunolocalization studies indicate that PRMT3 is cytoplasmic, while PRMT1 is localized in the nucleus. Although yeast two-hybrid data indicate that PRMT3 has the potential to form homo-oligomers as well as hetero-oligomers with PRMT1, gel filtration analysis indicates that PRMT3 is monomeric (Tang *et al.*, 1998). Crystallographic evidence, however, indicates that the PRMT3 core forms a homo-dimer when crystallized in the presence of a cofactor (*S*-adenosylhomocysteine). These homo-dimers, and perhaps even PRMT1/PRMT3 hetero-dimers, may also be present in equilibrium with the monomers *in vivo* (Zhang *et al.*, 2000), but are unable to survive the process of cell disruption and gel filtration analysis (Tang *et al.*, 1998), possibly due to the thermodynamics of the association/dissociation.

d. Protein arginine methyltransferase 4 / Coactivator-associated arginine methyltransferase 1

Protein arginine methyltransferase 4 (PRMT4), or Coactivator-associated arginine methyltransferase 1 (CARM1), is a protein that associates with the p160 family of coactivators, which are nuclear hormone regulated transcription factors. In the mouse, CARM1 is expressed as a 3.8-kb mRNA transcript in a wide variety of tissues. It shows 30 percent identity at the amino acid level to human PRMT1, and has been shown to preferentially methylate histone H3. CARM1 regulates transcription by

binding to the C-terminus of these p160 coactivators, in the AD2 activation domain. Mutations affecting the ability of CARM1 to bind S-adenosylmethionine have been shown to significantly reduce both its methyltransferase and coactivator activities (Chen *et al.*, 1999; Chen *et al.*, 2000; Stallcup *et al.*, 2000).

e. Protein arginine methyltransferase 5 / Janus kinase binding protein 1

Protein arginine methyltransferase 5 (PRMT5) is a cytoplasmic, homo-oligomeric protein that was originally called Janus kinase binding protein 1 (JBP1) (Frankel and Clarke, 2000; Rho *et al.*, 2001). The human JBP1 and its yeast homolog, Hsl7p, contain a methyltransferase motif and have been shown to exhibit protein arginine methyltransferase activity (Pollack *et al.*, 1999; Lee *et al.*, 2000). JBP1 binds to Janus kinase (Jak), which is involved in the Jak-Stat (signal transducers and activators of transcription) pathway (Pollack *et al.*, 1999). Yeast two-hybrid analysis indicates that it also interacts with the nonstructural protein 3 encoded by the hepatitis C virus (Rho *et al.*, 2001). Although the *in vivo* substrate for the methyltransferase activity of JBP1 has not been identified, a GST-PRMT5 fusion protein has been shown to methylate various Type I artificial substrates as well as MBP, a Type II substrate. It has not been determined, however, if the dimethylation pattern of MBP methylated by PRMT5 is symmetrical or asymmetrical; therefore, it is not clear if PRMT5 functions as a Type I, Type II, or a Type I/II enzyme (Rho *et al.*, 2001). The Jak-Stat pathway is involved in both signal transduction and cell cycle regulation; therefore, the involvement of a methyltransferase in this pathway suggests that protein arginine

methylation may play a role in cell signaling and/or cell cycle regulation (Pollack *et al.*, 1999).

B. Type II PRMTs

Since the discovery, 30 years ago, of the methylation of a specific arginine residue in the myelin basic protein (Baldwin and Carnegie, 1971b; Brostoff and Eylar, 1971) and the further identification of the modified residue as symmetrically dimethylated arginine (Diebler and Martenson, 1973), scientists have been searching for the enzyme responsible for this posttranslational modification. The enzymatic activity of the MBP-specific PRMT has been isolated (Crang and Jacobson, 1982; Sundarraj and Pfeiffer, 1973) and this enzyme activity is different from the activity of the histone-methylating PRMT (Kim *et al.*, 1997b). However, a single protein representing this activity has yet to be purified (Gary and Clarke, 1998).

The proposed kinetic mechanism for the enzyme involves AdoMet binding to the enzyme first, followed by rapid-equilibrium binding to the substrate (Young and Waickus, 1988). Studies of the Type II PRMT activity indicate that the preferred sequence for methylation is an arginine flanked on either side by glycines, although the enzyme retains at least partial activity when the N-terminal flanking glycine is substituted by certain amino acids (Hyun *et al.*, 2000). It was found that the MBP-specific PRMT requires a substrate of at least six residues for minimal activity, and that methyltransferase activity increases as the length of the substrate peptide increases (Ghosh *et al.*, 1991; Hyun *et al.*, 2000).

In addition to the standard inhibitors of all PRMTs, the MBP-specific PRMT is regulated by thyroid hormone and nitrous oxide (Amur *et al.*, 1984; Amur *et al.*, 1986). It is interesting to point out that animals exposed to nitrous oxide for prolonged periods of time show neurological effects due to demyelination (Amur *et al.*, 1986). In the mouse, it has been shown that the activity of this methyltransferase is synchronized with myelination (Crang and Jacobson, 1982; Kim *et al.*, 1988; Rawal *et al.*, 1991), and inhibition of the MBP-specific PRMT leads to incomplete myelination (Amur *et al.*, 1986).

C. Other PRMTs

RMT2 is a unique methyltransferase that was identified through analyzing SAM-binding sequence motifs found within the *Saccharomyces cerevisiae* genome. It is considered unique because it can methylate the δ-nitrogen of protein arginine residues (Niewmierzycka and Clarke, 1999), a modification that was only recently described (Zobel-Thropp *et al.*, 1998). RMT2 contains the conserved methyltransferase motifs I, post I, II, and III, and it shares significant identity to the human guanidinoacetate *N*-methyltransferase enzyme, which methylates the δ-nitrogen atom in guanidinoacetate. RMT2 homologs have been identified in *Schizosaccharomyces pombe* and *Arabidopsis thaliana*, but not in higher eukaryotes or prokaryotes. In yeast, RMT2 is essential for neither viability nor growth. Although the methyl-accepting substrate(s) for RMT2 has not been identified, it is believed that the substrates are fully methylated and that methylation occurs on the nascent proteins (Niewmierzycka and Clarke, 1999).

CHAPTER VII

SEQUENCE AND EXPRESSION ANALYSES OF CANDIDATE GENES

ABSTRACT

Sequence analysis of B179d14 identified two candidate genes for the *psrt* phenoypte: *Tip30/Cc3* and *Prmt3*. The *Tip30/Cc3* and *Prmt3* mRNAs are ubiquitously and constitutively expressed. Sequence analysis of the *Tip30/Cc3* cDNA indicates that the sequence is identical in the control and the two mutant strains, 723SJ and 1060SJ. There is evidence to suggest, however, that the *Prmt3* gene is altered in the mutants. Sequencing of cloned RT-PCR products failed to find a mutation in the 1060SJ *Prmt3* cDNA, but temperature-gradient capillary electrophoresis (TGCE) mediated heteroduplex analysis strongly indicates that a mutation is present in the 3' end of the transcript. Sequencing of the 723SJ *Prmt3* cDNA also failed to find a mutation; however, the 5' end of the 723SJ allele could not be amplified, either through standard RT-PCR or through 5' RACE. Southern blot analysis failed to verify a genomic rearrangement or deletion. However, TGCE heteroduplex analysis also indicates a mutation is present in the 3' end of the 723SJ transcript.

INTRODUCTION

In order to evaluate the candidacy of the two genes identified through sequence analysis of BAC B179d14, an expression analysis of each of the genes was performed in the mutant and control animals. The tissue-specific and temporal mRNA expression

patterns, as well as the cDNA sequence, will be useful in determining whether either of these genes could be responsible for the *psrt* phenotype.

MATERIALS AND METHODS

1. RNA Isolation

Total RNAs from brain were isolated using standard methods (Chirgwin *et al.*, 1979, Sambrook *et al.*, 1989). Poly $(A)^+$ RNAs were isolated from total RNAs using the Mini-Oligo (dT) Cellulose Spin Column Kit (5 Prime 3 Prime, Inc., Boulder, CO) following the manufacturer's instructions.

2. RT-PCR

Approximately 5μg of total RNA were reverse transcribed in a 20 μL reaction volume using Superscript II (Gibco BRL Life Technologies, Rockville, MD). A variety of primer combinations were used to amplify the cDNAs from both the *Prmt3* and the *Tip30/Cc3* genes (Table VII-1). Actin-specific primers (ORN 559/560) were used to amplify a 520 bp fragment of *β-actin* as the internal control (Alonso *et al.*, 1986).

3. Northern Blots

Total RNAs were run on a 1% denaturing formaldehyde agarose gel and electrophoresed in 1X MOPS buffer (Sambrook *et al.*, 1989). After electrophoresis, the gels were soaked for 30 minutes in 50 mM NaOH, rinsed in DEPC-treated water, neutralized two times for 15 minutes each in 0.1 M Tris-HCl, and blotted to Duralon

nylon membranes overnight in 20X SSC. After the transfer was complete, the blots were dried under vacuum at 80°C for two hours, and UV crosslinked at 250 mjoules/mm^2 in a Stratalinker (Stratagene, LaJolla, CA) before prehybridization.

4. RACE

5' RACE was performed using the Clontech SMART™ RACE Kit (Clontech, Inc.; Palo Alto, CA) according to the manufacturer's instructions. RACE was performed on first strand cDNAs generated from both poly $(A)^+$ RNAs and total RNAs from brain using both standard and touchdown PCR programs. Primers for the reaction were ORN 645 and ORN772, used with the universal primer and nested universal primer supplied with the RACE kit (Table VII-1).

5. MTC Panels

The mouse MTC (multiple tissue cDNA) panel (Clonetech, Inc.; Palo Alto, CA) was used to analyze, by PCR, expression of the candidate genes in different mouse tissues. This panel includes cDNAs from heart, brain, spleen, lung, liver, skeletal muscle, kidney, testis, and four different embryo stages (7-, 11-, 15-, and 17-day). The *Prmt3* PCRs utilized primers ORN 646 and 648, which amplify a 586 bp fragment from the 3' end of the transcript. The *Tip30/Cc3* PCRs utilized primers ORN 663 and 664, which amplify a 1060 bp fragment covering the entire coding sequence of the transcript. Actin-specific primers (ORN 559/560) were used to amplify a 520 bp fragment of *β-actin* as an internal control (Alonso *et al.*, 1986). For PCR, the conditions were 94° for

5 min, followed by 30 cycles of 94° for 30 sec, 1 min at the appropriate annealing temperature (60°C for ORN 663/664, 52°C for ORN 559/560, and at 62°C for ORN 646/648), and 72° for 3 min., then a final five minute extension at 72°C. Products were analyzed on a 1% agarose gel with ethidium bromide staining (Sambrook *et al.*, 1989).

6. Cloning and Sequencing RT-PCR Products

The PCR products obtained were cloned using the pGEMT-Easy II cloning kit (Promega Corporation, Madison, WI), following the manufacturer's instructions, and analyzed by fluorescent sequencing with the Big Dye Terminator Kit (PE Biosystems, Foster City, CA). The products were run on an ABI377 DNA sequencer and the sequence analyzed using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) to verify the identity of the cloned products.

7. Assembling cDNA Sequences

Individual sequencing runs from the RT-PCR products were assembled into contigs representing the entire cDNAs using the Sequencher 3.1 software (Gene Codes Corp., Ann Arbor, MI).

8. Mutation Screening by Temperature-Gradient Capillary Electrophoresis

RT-PCR amplified fragments of the *Prmt3* cDNA from the control (BJR) and the two mutant strains (723SJ and 1060SJ) were analyzed for mutations using

temperature-gradient capillary electrophoresis (TGCE) (Gao and Yeung, 2000). Six overlapping fragments were analyzed, ranging in length from 204 bp to 530 bp (Figure VII-2). The analysis was performed by SpectruMedix Corporation (State College, PA), in collaboration with Cymbeline T. Culiat of ORNL.

9. Genomic Southern Blotting

Genomic DNAs from 723SJ and 1060SJ *psrt* mutants and the BJR control were digested overnight with *BstX*I, *Eco*RI, *Pst*I, and *Taq*I (Gibco BRL Life Technologies, Rockville, MD) using the following reaction cocktail: 10μg of miniprep DNA, 4 μL of 10 X buffer (React2 for with *Pst*I and *Taq*I, and React3 for *BstX*I and *Eco*RI), 2 μL of enzyme, and sterile water to make a total volume of 41 μL. Digestions occurred overnight at 37°C except those using *Taq*I, which were digested at 65°C. Digested DNAs were loaded on a 0.8% agarose gel which was run in 1X TAE and Southern blotted using standard protocols (Sambrook *et al.*, 1989).

RESULTS

1. *Tip30/Cc3* cDNA Sequence

The *Tip30/Cc3* cDNA was amplified using primers ORN 663 and ORN 664, which amplify a 1060 bp fragment that includes the entire coding sequence of the *Tip30/Cc3* cDNA. Three sets of primers were also used to amplify smaller, overlapping segments of the cDNA: ORN 663 and ORN 728, which amplify a 615 bp segment in the 5' end of the cDNA; ORN 664 and ORN 727, which amplify a 464 bp segment in

the 3' end of the cDNA; and ORN 723 and ORN 726, which amplify a 560 bp segment in the middle of the cDNA that overlaps the other two segments (Figure VII-3).

The consensus sequence was determined by combining all sequencing data, including data from both the cloned RT-PCR fragments and from the direct sequencing of RT-PCR products. No point mutations or other discrepancies in the sequence were found between the control (BJR) strain and the two mutant strains (723SJ and 1060SJ).

2. *Tip30/Cc3* Expression Analysis

Primers ORN 663 and ORN 664 were used to look for expression in tissues from the Mouse MTC panel. A single 1060 bp fragment was detected for all twelve cDNAs tested (Figure VII-4), indicating that *Tip30/Cc3* is ubiquitously expressed in mouse adult and embryonic tissues.

3. *Prmt3* cDNA Sequence

The *Prmt3* cDNA was amplified from the two mutant strains and the control using two sets of primers that amplify overlapping segments of the cDNA: ORN 646 and ORN 648, which amplifies a 586 bp segment in the 3' end of the cDNA that includes the stop codon and part of the 3' UTR; and ORN 644 and ORN 647, which amplify a 685 bp segment in the middle of the cDNA that overlaps the other two segments (Figure VII-5). No amplification product could be obtained from the 723SJ *psrt* cDNA using ORN 645 and ORN 704, which amplify a 569 bp segment in the 5' end of the cDNA that includes the start codon and a short segment of the 5' untranslated

region (UTR). There was, however, an amplification product from the 1060SJ mutant and the BJR control cDNAs.

When visualized on an ethidium bromide stained agarose gel, the ORN 644/647 product amplified from 1060SJ mutants appears to be a single product; however, during the process of cloning and sequencing, a single clone was isolated that carries an alternatively spliced insert. This anomalous clone is missing bp 844 to 1063, which corresponds to exons nine and ten. Therefore, although this transcript is in low abundance, at least one alternatively spliced transcript exists for the *Prmt3* gene in the 1060SJ animals.

The *Prmt3* consensus sequence was determined by combining all sequencing data, including data from both the cloned RT-PCR fragments as well as from the direct sequencing of RT-PCR products. No point mutations or other discrepancies in the sequence were found between the control (BJR) strain and the two mutant strains (723SJ and 1060SJ).

4. 5' RACE

RACE was utilized in an attempt to amplify the 5' end of the 723SJ *Prmt3* cDNA, which could not be amplified by standard RT-PCR. The RACE controls were successfully amplified; however, repeated attempts at amplification of the 723SJ cDNA using both total RNAs and poly $(A)^+$ RNAs were unsuccessful. These results suggest the presence of a possible genomic rearrangement or deletion in the 723SJ *Prmt3* allele.

5. *Prmt3* Expression Analysis

Primers ORN 646 and ORN 648 were used to look for expression in tissues from the Mouse MTC panel. A 586 bp fragment was detected for all twelve cDNAs tested, and a second, larger fragment was also detected in skeletal muscle and possibly in heart (Figure VII-6). This indicates that *Prmt3* is ubiquitously expressed in mouse adult and embryonic tissues, and that at least one larger, alternatively spliced transcript is present in skeletal muscle and heart.

Northern blot analysis of brain total RNA probed with a *Prmt3* EST (Genbank Accession #ai156519) indicates that the *Prmt3* transcript is approximately 2.4-kb, which agrees with the size previously reported for rat brain (Tang *et al.*, 1998). The 723SJ *psrt* transcript appears to be slightly shorter than the 1060SJ *psrt* transcript and the controls (Figure VII-7). However, this difference is slight, at best, and has not been quantified.

6. Detection of Genomic Rearrangement or Deletion

The genomic Southern blot of DNAs from the control strain (BJR) and the two mutant strains (723SJ and 1060SJ) was probed with RT 645/704, a RT-PCR product from the 5' end of the *Prmt3* cDNA transcript. The enzymes used to digest the genomic DNA were chosen because the genomic sequence (obtained from the BAC containing the *Prmt3* and *Tip30/Cc3* genes, B179d14) indicates the presence of restriction sites for each of the enzymes within the first 20-kb of the *Prmt3* gene and the 5-kb of genomic sequence directly upstream of the gene. In addition, the restriction map of all four

enzymes shows an abundance of fragments ranging from 0.3-kb to 6-kb, which could be detected on a Southern blot (Table VII-8). If either a genomic rearrangement or deletion were present in one of the mutant strains, then the size of the band(s) might be altered relative to the other strains. The autoradiograph of the probed blot, however, showed identical bands for the three DNAs, each digested with four different enzymes. This indicates that, for the four enzymes used, no deletion or rearrangement is detected (Figure VII-9).

7. *Prmt3* TGCE Mutation Analysis

A TGCE analysis is performed by mixing the amplified cDNA from a homozygous (or hemizygous) mutant with that from the control animal. The samples are then denatured and allowed to reanneal. If both the mutant and control strands are identical, then a homoduplex is formed and the electropherogram shows a single, smooth peak. However, if mutations are present, including single nucleotide polymorphisms (SNPs), then both heteroduplexes and homoduplexes will form when the samples reanneal. Since heteroduplexes and homoduplexes have different electrophoretic mobilities in the electrophoresis matrix [in this case, a poly(vinylpyrrolidone) solution], the shape of the resulting electropherogram will no longer be a single, smooth peak. Rather, it can range from a single peak with a "shoulder" to a multiplet of up to four individual peaks representing two complementing homoduplexes and both possible heteroduplexes.

The TGCE analysis performed by the SpectruMedix Corporation indicates that mutations are present in both of the mutant strains. Heteroduplex formation occurred in the 1060SJ cDNA amplified with ORN 648/676, which is from the 3' end of the cDNA (Figures VII-10 and VII-11). The overlapping segment, ORN 646/677, does not form heteroduplexes; therefore, there appears to be a mutation somewhere in the 162 bp of the fragment that does not overlap with ORN 646/677. The heteroduplex was detected in both of the two independently generated cDNAs from 1060SJ mutants that were analyzed.

In addition, heteroduplex formation was also detected in the 723SJ cDNA in two overlapping segments, ORN 646/677 and ORN 648/676 (Figures VII-10 and VII-11). These two segments represent the 3' end of the cDNA; however, since both segments form heteroduplexes, the mutation is most probably located in the 89 bp that form the overlap between the two segments. The heteroduplex was detected in a single 723SJ cDNA (rather than in both cDNAs) because one of the two submitted samples failed to produce a signal. The single 723SJ cDNA did, however, show heteroduplex formation with both of the control samples.

DISCUSSION

The *Tip30/Cc3* gene appears to be intact and the sequence is identical in the control strain and in *psrt* homozygotes animals from the two mutant strains. Although a sequence analysis using a cDNA-based primer set indicates that the *Prmt3* cDNA sequence is unchanged between the control and the two mutant strains, there is evidence

to suggest that the *Prmt3* gene is altered in the mutants. Northern blot analysis suggests a possibly shortened transcript in the 723SJ mutants, and the 5' end of the 723SJ transcript could not be amplified, either by standard RT-PCR or by RACE. A genomic rearrangement or deletion could be the cause of the problem; however, Southern blot data failed to verify a genomic alteration. TGCE data strongly indicate that mutations are present in the *Prmt3* transcripts from each of the two mutant strains, both in the 3' end of the transcripts, with the 723SJ mutation occurring in a position more 5' than the 1060SJ mutation. Additional sequence analysis will be required to verify these mutations as well as to ascertain conclusively if there is an aberration in the 5' end of the 723SJ transcript.

CHAPTER VIII

ANALYSIS OF METHYLARGININE IN THE MYELIN BASIC PROTEIN OF *psrt* MUTANT ANIMALS

ABSTRACT

To assess the methylation status of the arginine-107 (R-107) residue of the myelin basic protein (MBP), MBP was extracted from *psrt* mutant and non-mutant mouse brains. The MBP was hydrolyzed and the amino acids derivatized with *N*-methyl-*N*-(*tert*-butyldimethylsilyl)trifluoroacetamide (MTBSTFA) to produce *tert*-butyldimethylsilyl (*t*-BDMS) derivatives. The derivatized amino acids were analyzed by gas chromatography with mass spectrometric detection (GC/MS) to determine the methylation status of the arginine residues in MBP. The hydrolysate from the control animals contained the expected methylarginine residue at quantifiable levels, while the hydrolysate from the *psrt* mutants contained no detectable methylarginine.

INTRODUCTION

One of the candidate genes for the *psrt* phenotype is a protein arginine *N*-methyltransferase, *Prmt3*, whose specific *in vivo* substrate is unknown. Although there are several proteins known to contain methylarginine residues, the myelin basic protein (MBP) is the only known methylarginine (Me-R)-containing protein that is involved in neurological function. MBP contains a single Me-R residue, arginine-107 (R-107), that can be either monomethylated or symmetrically dimethylated.

If *Prmt3* is the gene responsible for *psrt*, then the mutants will exhibit abnormal protein arginine methylation. Furthermore, if MBP is the substrate for *Prmt3*, then abnormal methylation patterns would be present in the MBP isolated from the mutants. To investigate this possibility, MBP from the brains of the *psrt* mutants and their non-*psrt* littermate controls were analyzed for the presence of methylargine (Me-R).

MATERIALS AND METHODS

1. Isolation of MBP

MBP was extracted from two groups of 14 day old brains: seven pooled *psrt* brains, totaling 2.08 grams, and seven pooled control brains, totaling 2.61 grams. The pooled mutant sample class included five 1060SJ *psrt* homozygotes (genotype *ru2 m p / ru2 m p*) and two hemizygotes (genotype *ru2 m p / Df[ru2-p]*); the pooled 1060SJ control animals included five wild-type animals (genotype *ru2* + + / + + *p*) and two non-carrier *ru2* animals (genotype *ru2* + + / *ru2* + +). MBP was isolated using a variation of the technique developed by Maatta *et al.* (1997). Briefly, the brains are homogenized in chloroform (99.9%, spectrophotometric grade, ACS reagent; Burdick and Jackson Laboratories, Inc., Muskegon, MI) and the homogenate is centrifuged to isolate the MBP-containing organic phase from the pellet and the aqueous phase. The pellet is re-homogenized in chloroform and the process is repeated. The organic phases are combined, washed in water ('Baker Analyzed'® HPLC reagent; J.T. Baker, Phillipsburg, NJ) and centrifuged again. The MBP-containing organic phase is removed, mixed with ('Baker Analyzed'® HPLC reagent) methanol (J.T. Baker,

Phillipsburg, NJ) and vortexed. Acidified water (containing HCl) is added, and the mixture is vortexed again. At this stage, the MBP is dissolved in the acidic aqueous phase, which is removed and dried in a rotary evaporator.

2. Hydrolysis and Derivatization of MBP

The dehydrated MBP, which is a filmy residue in the bottom of the round bottom flask, is dissolved in 500 μL of 6M constant-boiling HCL (Pierce Chemical Co., Rockford, IL) and transferred to a hydrolysis tube. The hydrolysis is carried out overnight at 110°C to 120°C under vacuum. The hydrolysate is dried by warming to 55°C under a stream of nitrogen, then derivatized at 80°C for one hour using 30 μL of MTBSTFA + 1% TBDMCS (*tert*-butyldimethylchlorosilane, a catalyst) (Pierce Chemical Co., Rockford, IL) and 30 μL of pyridine (99.8%, anhydrous; Sigma, Inc., St. Louis, MO).

3. GC/MS Analysis

Samples were analyzed using a Hewlett-Packard model 5890 gas chromatograph with a HP-5MS capillary column (30m x 0.25mm i.d.; 0.25 μm film thickness). The detector was a model 5989A quadrupole mass spectrometer operating in electron impact ionization mode. The ion source was maintained at 70 eV and 200°C. The carrier gas was helium and the column head pressure was 7 psi. The injector temperature was 250°C and the temperature program was as follows: 60°C for two minutes, then ramp

at 5°C per minute for 44 minutes to 280°C and hold for five minutes. The total run lasts 51 minutes.

4. Amino Acid Standards

Glycine (gly), lysine (lys), methylarginine (me-R), and tyrosine (tyr) (Sigma, Inc., St. Louis, MO) were individually dissolved in methanol (me-R), (1:1, v:v) methanol:water (gly, lys, and phe) and pH 2 water (tyr). Aliquots from these solutions were combined and dried under nitrogen at 55°C then derivatized as described above to produce solutions containing 50, 100, 200, and 300 ng/μL of each amino acid. One μL of each of these combined standards was injected and analyzed using capillary GC/MS as described above. The peak areas from the chromatographs were utilized to produce a calibration curve for each amino acid, which was used to quantitate the amount of these amino acids present in the MBP samples from the mutant and control animals.

RESULTS

1. Calibration Curves for Amino Acid Standards

The calibration curves for the four amino acid standards (glycine, lysine, methylarginine, and tyrosine) are all roughly linear over the concentration ranges tested (50 ng to 300 ng injected) (Figure VIII-1). Extrapolation from the curves for each of the standards allows for quantitation of the standards in the samples (Table VIII-2).

2. Amino Acid Analysis of Chromatographs of Mutant and Non-Mutant Mice

The total ion chromatograph of the derivatized hydrolysate of the control mice shows a methylarginine peak area of 3,523,593 which corresponds to approximately 100 ng of methylarginine. The total ion chromatograph of *psrt* mutant brains does not have a methylarginine peak, indicating that the level of Me-R present is below the detection limit (Table VIII-2).

The chromatographs of the derivatized MBP hydrolysate from the *psrt* mutant animals and the non-mutant controls are similar, with many peaks in common. They are not, however, exactly alike. The peaks on the chromatograph from the mutants tend to be smaller than those from the controls, and there are peaks from the control chromatograph that are missing in the mutants. The Me-R peak is clearly seen in the control chromatograph, but is missing in the *psrt* chromatograph (Figures VIII-3 and VIII-4).

DISCUSSION

The GC/MS analysis indicates that there are no detectable levels of Me-R present in the MBP of the pooled *psrt* mutant brains. This could be due to several different reasons. Although the same number of brains were pooled for the *psrt* and control samples, the *psrt* brains are smaller; therefore, the total mass of mutant brains is 20.3% less (2.08 grams) than the total mass of the control brains (2.61 grams). In addition, the *psrt* mutants are runted and appear to be developmentally delayed. Even

though MBP mRNA and protein expression have been occurring since before day eight in normal mice, it is possible that there is a delay in this process in the mutants, which would mean that not as much MBP and, therefore, Me-R is present in the mutants as is present in age-matched controls. The quantity of material present in the mutant MBP hydrolysate as compared to the control MBP hydrolysate (as determined by comparing peak areas of tyrosine and lysine to the calibration curves) tends to support this notion (Table VII-2). The fact that there was a visible film on the flasks at the end of the MBP isolation and the presence of quantifiable amounts of amino acids in the mutant chromatographs, however, implies that protein was indeed present. This assay is designed to isolate relatively pure MBP and the pH is adjusted to below two in order to insure that other basic proteins, such as histones, will precipitate out of solution, leaving only the MBP in the final acidic aqueous phase (Maatta *et al.*, 1997). Therefore, MBP appears to be present in the mutant brains, even though Me-R was not detected.

This evidence suggests that a defect in protein arginine methylation is present in the mutants; however, it is not conclusive. There was less MBP present in the mutants than in the control brains, and the quantity of Me-R detected in the control brains was nearing the detection limit for the method. Since multiple brains are required for this assay and only a few mutant brains were available for analysis, a single trial was conducted using only one of the two mutant strains (1060SJ). A more complete analysis involving multiple samples from both mutant strains is needed to conclusively determine if the biochemical defect in *psrt* mutants affects the methylation of the R-107 residue of MBP.

CHAPTER IX

EVALUATION OF CANDIDATE GENES

INTRODUCTION

A positional cloning strategy, consisting of physical mapping followed by shotgun sequencing of genomic clones, was employed to identify two genes (*Tip30/Cc3* and *Prmt3*) mapping into the same *p*-region deletion interval as the *psrt* phenotype. In addition, a positional candidate strategy was used to identify a third gene, *Glyt2*, mapping into the homologous region of the human genome. What follows is a discussion of each of the three genes and an assessment of their candidacy.

DISCUSSION

1. Assessment of the Candidacy of *Tip30/Cc3*

The *Tip30/Cc3* gene was identified as a candidate gene because it maps into the same *p*-region deletion interval as the *psrt* phenotype. It was discovered through sequence analysis of a BAC clone that is part of the physical map of the region. Since *Tip30/Cc3* is involved in transcription, a mutation in this gene would affect the transcriptional machinery and/or processes in the mutant animals. A ubiquitously expressed, transcription-related gene such as *Tip30/Cc3* might be responsible for a runting phenotype, especially since Cc3 is involved in promoting apoptosis; however, Tip30/Cc3 has not been linked to neurological function.

The entire coding sequence of the *Tip30/Cc3* cDNA was sequenced in the control (BJR) and both mutant strains, and no mutations were found. Based on this information and the inability to link the function of the *Tip30/Cc3* gene product to a neurological phenotype, it is not considered a strong/viable candidate for the *psrt* phenotype.

2. Assessment of the Candidacy of *GlyT2*

The human genome sequencing databases were searched for candidate genes mapping within 11p14-p15, the region homologous to the *psrt* region of Mmu 7. The aim was to find genes associated with phenotypes similar to *psrt*, namely runting, seizures, and prenatal or juvenile lethality. Although no known human disease genes were found to map within the interval, the *glycine transporter type 2* (*GLYT2*) gene was identified as a candidate because of its involvement in neurological function.

There are several reasons why *Glyt2* could be the gene responsible for the *psrt* phenotype. A mutation in *Glyt2* could alter glycine transport in the mutant animals, which would lead to nervous system dysfunction and neural degeneration. The dysregulation of neurotransmitter transport alters the amount of neurotransmitter present in the cell, in both the cytoplasm and in the synaptic vesicles. This has been shown to alter synaptic strength (Liu and Edwards, 1997; Pothos *et al.*, 2000). Neurotransmitter toxicity, resulting from the accumulation of neurotransmitter in the cell, can lead to neuropathological conditions. Dopamine toxicity, for example, has been associated with Parkinson's disease (Liu and Edwards, 1997), and elevated

extracellular concentrations of glutamate have been associated with a number of neurodegenerative diseases, including ALS (Lou Gehrig's disease), Huntington's disease, and Alzheimer's disease (Gegelashvili and Schousboe, 1997).

The seizure phenotype seen in *psrt* animals is compatible with a loss of inhibitory neurotransmission. Their limbs are stiff and extended for the duration of the seizure, not unlike the effects of strychnine poisoning, which causes convulsions and muscle contractions and is often fatal (Alberts *et al.*, 1996). Strychnine is a glycine antagonist, meaning that it competes with glycine at the glycinergic synapses and interferes with inhibitory neurotransmission. If the *psrt* phenotype were caused by mutations in *Glyt2*, then glycine transport would be dysfunctional. Altered glycine transport would cause the accumulation of glycine at the postsynaptic membrane, which could have a neurotoxic effect on the nervous system and lead to the death of the animal. The affect of the accumulation of glycine at the postsynaptic membrane is unknown; however, it is likely that excess glycine would cause an increase in inhibitory neurotransmission, which is the opposite effect to the observed phenotype.

In the mouse, CNS and PNS myelination occur after birth (Raine, 1984a). The onset of seizures in the *psrt* mutants occurs between seven and ten days of age, just as the nervous systems are developing. Glycinergic motoneurons are associated with rhythmic movements, such as locomotion and respiration. These movements involve repeating cycles of precisely coordinated excitation and inhibition (Rekling *et al.*, 2000). Therefore, if glycine transport were dysfunctional in the *psrt* mutants, there would be an effect on rhythmic movements, and the visible effects might increase as the

nervous system matures and rhythmic movements increase. This appears to be the case with *psrt* mutants, since the onset of seizure activity occurs at about seven to ten days after birth.

The *psrt* phenotype could be caused by a mutation in *Glyt2*. Such a mutation would affect glycine transport, and altered glycine transport would severely affect nervous system function. The seizures in the *psrt* mutants are first observed during the time that the nervous systems are developing, and death occurs within a week of the onset of seizures. Therefore, there is a temporal correlation between the observed *psrt* seizure phenotype and the function of the *Glyt2* protein. Since *Glyt2* was discovered near the end of this project when the molecular work was nearly complete, no molecular data are available for it. Molecular expression data are necessary to thoroughly evaluate the candidacy of *glyt2*; without them, no conclusions can be drawn for the candidacy of *glyt2*. Therefore, although *glyt2* appears to be a viable candidate for the *psrt* phenotype, a full assessment of its candidacy must, by necessity, await molecular expression data and sequencing.

3. Assessment of the Candidacy of *Prmt3*

In order to assess the candidacy of *Prmt3* as the gene responsible for the *psrt* phenotype, it is necessary to examine the link between protein arginine methylation and neurological function. The only molecularly characterized mammalian mutation involving a protein arginine methyltransferase gene involves the mouse *Prmt1* gene. Unfortunately, *Prmt1* null homozygotes die at around day E6.5; therefore, it is

impossible to determine the effect(s) of *Prmt1* on the developing nervous system. There is, however, one other well-known link between protein arginine methylation and nervous system function, and it involves the myelin basic protein (MBP).

Myelin is a membrane architecture found within the central (CNS) and peripheral (PNS) nervous systems. Its function is to insulate the axon, thereby increasing the velocity of the nerve impulse (Kim *et al.*, 1997b; Raine, 1984a; Alberts *et al.*, 1996). Myelin contains about 80% lipid and 20% protein, and MBP comprises approximately 30% of the total CNS protein (Lees and Brostoff, 1984).

MBP is a peripheral membrane protein localized to the major dense line regions of compact CNS and PNS myelin. It is located on the cytoplasmic side of the phospholipid bilayer of either the oligodendrocyte in the CNS or the Schwann cell in the PNS (Raine, 1984a). During CNS development, certain hnRNPs bind to and transport MBP mRNA from the nucleus to the distal ends of the oligodendrocyte processes. Once there, the MBP mRNA is anchored to the myelin membrane and translated into protein (Hoek *et al.*, 1998). As myelinogenesis progresses, the cytoplasm is extruded and the membranes begin to wrap spirally until multiple layers of membrane and protein surround each axon (Omlin *et al.*, 1982; Raine, 1984a). A similar process occurs in the Schwann cell in the PNS, except that MBP comprises a smaller percentage of the total protein in PNS myelin (Lees and Brostoff, 1984; Raine, 1984a).

MBP undergoes several posttranslational modifications; however, the methylation of the arginine-107 (R-107) residue appears to be a nearly universal

modification. Only one exception, carp, has ever been noted (Kies *et al.*, 1972; Deibler and Martenson, 1973; Young *et al.*, 1987). The methylation of the R-107 residue is critical to the function of MBP within the myelin membrane, affecting the ability of the protein to bind to the myelin membrane. This methylation is also involved in maintaining the integrity of the myelin sheath (Kim *et al.*, 1997b; Zand *et al.*, 1998; Boggs *et al.*, 1999), and has been implicated in multiple sclerosis (MS), possibly as the autoantigen responsible for the disease (Moscarello *et al.*, 1994). Irrespective of the cause of the disease, however, the MBP from MS patients contains less methylarginine than would be expected (Moscarello *et al.*, 1994; Rawal et al., 1995a; Kim *et al.*, 1997b). Therefore, there appears to be a correlation between MBP methylation and the demyelination that occurs in MS, which supports the assertion that a defect in MBP methylation could be responsible for the neurological phenotype seen in *psrt* mutants.

There are at least eight known hereditary diseases that cause a dysmyelination phenotype in humans (metachromatic leukodystrophy, Krabbe's disease, adrenoleukodystrophy, Refsum's disease, Pelizaeus-Merzbacher disease, Alexander's disease, Canavan's disease, and phenylketonurea). However, none of these diseases is known to directly involve MBP (Raine, 1984b). In the mouse, there are several well-characterized mutants with a dysmyelination phenotype. These include jimpy (*jp*), quaking (*qk*), murine muscular dystrophy (*dy*), Trembler (*Tr*), shiverer (*shi*) and myelin deficient (shi^{mld}) (Raine, 1984b). Although MBP levels are lowered in some of these mutants (Bourre *et al.*, 1980; Jacque *et al.*, 1983), only two of them, *shi* and shi^{mld} have

been shown to be mutations in the MBP gene itself (Roach *et al.*, 1983; Popko *et al.*, 1987).

The *shiverer* (*shi*) mutation is an autosomal recessive mutation that maps to Mmu 18 (Barbarese *et al.*, 1983; Sidman *et al.*, 1985; Raine, 1984b). It is a lethal mutation, characterized by seizures, with death occurring between two and five months of age (Readhead *et al.*, 1987). *Shiverer* mice exhibit severe dysmyelination. The CNS in these mice is characterized by grossly disorganized myelin sheaths (Rosenbluth, 1980a); however, only subtle changes are detectable in the PNS myelin (Kirschner and Ganser, 1980; Rosenbluth, 1980b). Brains of *shi* mice contain virtually no MBP mRNA or protein (Roach *et al.*, 1983). The *myelin deficient* (shi^{mld}) mutation is allelic to *shi*, with shi^{mld} animals exhibiting a similar, but less severe phenotype. The CNS of shi^{mld} mice is hypomyelinated, and death occurs at a later time, between five and nine months of age. Unlike *shi* mice, however, shi^{mld} homozygotes express reduced, but detectable, levels of MBP mRNA and protein (Popko *et al.*, 1987). Molecular analysis indicates that the *shi* mutation involves a nearly complete deletion of the MBP gene (Roach *et al.*, 1983). The shi^{mld} mutation, on the other hand, stems from the presence of multiple, linked copies of the MBP gene, with some of these copies containing a rearrangement, but at least one intact MBP gene (Popko *et al.*, 1987).

If the function of the PRMT3 protein is to methylate the R-107 residue of MBP, then it becomes necessary to address the question of why the phenotype of *psrt* animals is so much more severe than the phenotypes of either *shi* or shi^{mld} mice. Since the *shi* allele is a deletion of MBP and the *shi* phenotype is characterized by the near absence of

MBP, *shi* is considered to be a null mutation; yet the animals live between two and five months. The *shi*^{mld} mutation is hypomorphic, with very low levels of MBP detected in the mutants, and these animals live between five and nine months. Conversely, *psrt* animals, which presumably have a wild-type MBP gene but may lack the ability to properly methylate the MBP protein, live only 15 to 18 days after birth.

Although this evidence appears at first to be to the contrary, it does not preclude the possibility that the function of the PRMT3 protein is to methylate the MBP. The *shi* mutation is carried on a Swiss Vancouver (SWV) background (Bird *et al.*, 1978; Hogan and Greenfield, 1984), while *psrt* is on a BJR background. The genetic background of the *shi* animals is different from the *psrt* animals, and it has been shown that genetic background can have a dramatic effect on phenotype. For instance, mice homozygous for a transgenically inserted null allele of the epidermal growth factor receptor (*Efgr*) display a variety of phenotypes, depending on the genetic background of the transgenic animal. These phenotypes are all lethal; however, death can occur anytime from peri-implantation (on a CF-1 background) to mid gestation (on a 129/Sv background) to 20 days after birth (on a 129/Sv x C57BL/6 x MF1 background) (Sibilia and Wagner, 1995; Threadgill *et al.*, 1995). Therefore, mutations affecting the same biochemical processes (i.e. myelination) in different strains of mice (i.e. SWV and BJR) can have widely differing phenotypes due to the differences in genetic background.

Since the *psrt* phenotype includes runting in addition to the neurological (seizure) phenotype, it is also possible that PRMT3 may have more than one *in vivo* substrate and, consequently, a mutation in *Prmt3* could cause more than one effect.

This idea of multiple substrates is supported by the fact that the MBP gene is expressed only in the nervous system, in oligodendrocytes and Schwann cells (Raine, 1984a), while *Prmt3* is ubiquitously expressed. Therefore, the differences in phenotype between *psrt* and *shi* animals could be caused by either one or both of two things: (1) the mutations are on different genetic backgrounds, and (2) there is more than one substrate for PRMT3, which could amplify the pheyotypic effect of a mutation in the gene.

PRMT3, unlike other known PRMTs, contains a zinc-finger motif, and it has been suggested that this specific part of the protein is responsible, at least in part, for its substrate specificity (Frankel and Clarke, 2000). Zinc fingers are associated with ligand binding, most commonly with protein-DNA binding, but also with protein-RNA and protein-protein interactions (Mackay and Crossley, 1998). Since the MBP mRNA is bound to an hnRNP, it seems reasonable that the enzyme responsible for methylating the MBP might be capable of binding either to the MBP mRNA, the hnRNP that carries it to its final destination, or the MBP itself. Furthermore, immunolocalization studies indicate that PRMT3 is primarily cytoplasmic (Tang *et al.*, 1998). Once transcription is complete (in the nucleus), the nascent MBP mRNA binds to an hnRNP and is transported from the nucleus to the end of the oligodendrocyte process, where it is anchored to the membrane and translated into protein (Hoek *et al.*, 1998). It is reasonable to expect that any protein responsible for modifying MBP would be localized in the cytoplasm, since the MBP mRNA spends most of its lifetime in the cytoplasm and the MBP protein is entirely cytosolic. Therefore, the ligand-binding

capabilities and the intracellular location of the PRMT3 protein are consistent with those of an enzyme that could modify MBP.

Earlier studies have suggested that a single methyltransferase could be responsible for all three types of methylarginine; however, it was later determined that the Type I activity (producing Me_2Arg and Me_2(asym)Arg) and the Type II activity (producing Me_2Arg and Me_2(sym)Arg) could be separated (Lee *et al.*, 1977). Despite the isolation of specific methyltransferase activities and numerous efforts to purify the proteins responsible, all attempts to purify to homogeneity and sequence the proteins responsible for these activities have failed (Lin *et al.*, 1996; Gary and Clarke, 1998; Tang *et al.*, 2000). It is interesting to note that the enzyme responsible for the Type I activity (producing solely MeArg and Me_2(asym)Arg) is inhibited by MBP, a substrate for the Type II activity (Park *et al.*, 1989). Competitive inhibition of an enzyme by its product is a common phenomenon; however, MBP from mature myelin contains only MeArg and Me_2(sym)Arg, making it a product of the Type II, and not the Type I PRMT enzyme. This suggests that the criteria for classifying PRMT enzymes may not be absolute, and that MBP could be methylated by an enzyme that has been classified as Type I.

Although a GST-PRMT3 fusion protein has been shown to exhibit Type I PRMT activity, meaning that it produces solely MeArg and Me_2(asym)Arg residues in an artificial substrate (Tang *et al.*, 1998; Zhang *et al.*, 2000), there is evidence suggesting that native PRMTs and their GST fusion proteins may have different properties (Frankel and Clarke, 2000). Since the natural substrate for PRMT3 has not

been identified (Smith *et al.*, 1999) and the native protein has not yet been isolated and purified, it is impossible to know for certain how PRMT3 functions *in vivo*. Therefore, it is possible that the native PRMT3 protein, unlike the fusion protein containing only a part of the total enzyme, might exhibit some Type II PRMT activity *in vivo*.

Along these same lines, it has been assumed that the PRMT enzyme (Type I or Type II) is solely responsible for determining the methylation state of its substrate. There is evidence, however, to suggest that the context of, or the environment surrounding, an arginine residue can also affect its methylation state (Rawal et al., 1995a; Brahms *et al.*, 2000). In order for Sm proteins D1 and D3 to be properly (symmetrically) dimethylated, they must first be assembled into their native hetero-oligomeric complex. Unassembled D1 subunits are improperly (asymmetrically) dimethylated (Brahms *et al.*, 2000). It has also been shown that immature myelin contains a higher abundance of exon-2-containing MBP isoforms (Kruger *et al.*, 1999). These proteins would vary in their N-terminal residues, which are a part of the NAR (N-terminal acidic region) containing the methylated R-107 residue. Since the methylation status of MBP changes as the myelin matures (Rawal et al., 1995a), it is possible that the change in the environment near the R-107 residue is affecting the methylation status of R-107. Furthermore, it has been shown that MBP undergoes a steric transformation upon binding to lipids, with an increase in secondary structure of the protein (Lees and Brostoff, 1984). A change in environment could result either from the presence of different MBP isoforms at different developmental stages or through steric changes in the protein resulting from the association of MBP with other proteins or membrane

lipids. This opens the possibility that the enzyme responsible for modifying the MBP might act as a Type I enzyme under certain circumstances but, when the context of the R-107 changes, can act as a Type II enzyme. Therefore, despite reports that GST-PRMT3 exhibits no Type II activity, until the native protein and its substrate can be studied *in vivo*, it is at least a possibility that MBP could be an *in vivo* substrate of native PRMT3.

Perhaps the most compelling evidence supporting the candidacy of *Prmt3* in the *psrt* phenotype is the temporal correlation between myelination, PRMT activity, and the *psrt* phenotype. In mice exhibiting the *psrt* phenotype, the onset of seizures occurs approximately 7-10 days after birth, with seizure activity increasing in duration and intensity until the animal eventually dies, typically between 15 and 18 days after birth. In the mouse, MBP mRNA expression is first detectable by northern blot at around 7 days (Mathisen *et al.*, 1993; Haas *et al.*, 1995). MBP mRNA expression, MBP protein synthesis, MBP-specific methylase activity, and myelination occur synchronously, peaking at 16-18 days after birth (Campagnoni *et al.*, 1978; Carson *et al.*, 1983; Chanderkar *et al.*, 1986; Rawal *et al.*, 1991). Therefore, the onset of seizure activity in the *psrt* animals correlates with the onset of MBP expression and CNS myelination, and death occurs at the peak of myelinogenesis.

There are several reasons why *Prmt3* could be responsible for methylating the R-107 residue of MBP, a posttranslational modification that is critical to the function of MBP within the myelin membrane. If the PRMT3 protein is responsible for this methylation, then a mutation in the *Prmt3* gene could cause a lethal neurological

phenotype such as *psrt*. Since the *in vivo* substrates of PRMT3 are unknown, it is also possible that there are other PRMT3 substrates (such as bFGF) that, when improperly methylated, could lead to either a runting or a neurological phenotype.

CONCLUSIONS

Of the three genes examined in this work, *Prmt3* is the leading candidate for the cause of the *psrt* phenotype. The entire coding sequence of the *Tip30/Cc3* cDNA was sequenced for the control and both mutants, and no mutations were found. Based on this information and the inability to link the function of the *Tip30/Cc3* gene product to a neurological phenotype, it is not considered a strong candidate for the *psrt* phenotype. Although *Glyt2* does appear to be a viable candidate for a neurological phenotype, it was discovered only recently; therefore, a thorough assessment of its candidacy, including a molecular expression analysis, has not been completed.

The most probable assertion, based on the evidence presented, is that *Prmt3* is the gene responsible for the *psrt* phenotype. Protein arginine methylation has been linked to neurological function. The methylation of the R-107 residue of MBP is critical to the function of MBP within the myelin membrane. It has been shown that myelin is severely affected by either the absence of MBP (as in *shi* mice) or incorrectly methylated MBP (as in MS). The cellular localization pattern of the PRMT3 protein is the same as MBP. Furthermore, there is a temporal correlation between myelination and the *psrt* phenotype. Therefore, if the PRMT3 protein is responsible for methylating

the R-107 residue of MBP, then a mutation in the *Prmt3* gene could effect a lethal neurological phenotype such as *psrt*.

It is important to emphasize that PRMT3 could be responsible for the *psrt* phenotype even if MBP is not an *in vivo* substrate for the enzyme. Mutations in *Prmt3* would lead to the improper methylation of arginine residue(s) in its substrate(s). Since the *in vivo* substrates of PRMT3 are unknown, it is possible that other substrates (such as bFGF) exist that, when improperly methylated, could lead to either a runting or a neurological phenotype.

The molecular evidence presented in this work suggests that the *psrt* mutants have altered *Prmt3* alleles. Although sequencing efforts have failed to detect a mutation in the coding region of the cDNAs from either of the two homozygous mutants, preliminary data from a mutation screen utilizing temperature-gradient capillary electrophoresis (TGCE) (SpectruMedix, State College, PA) strongly suggest that a mutation is present in the 3' end of the transcript from the 1060SJ allele. In addition, MBP from 14 day old 1060SJ mutants contains no detectable amounts of MeArg, while MeArg can be easily seen in age-matched littermate controls. This indicates that there is a lack of MBP-specific PRMT activity in the 1060SJ mutants.

The TGCE data also suggest a mutation in the 3' end of the transcript of the 723SJ allele, which has been difficult to characterize molecularly. The 5'end of the 723SJ *Prmt3* transcript could not be amplified, either by standard RT-PCR or 5' RACE. Northern blot analysis indicates that the 723SJ transcript may be shorter than the 1060SJ allele and the control; however, this difference is only very slight, at best. This

suggests that some type of genomic rearrangement or microdeletion may be present in the 723SJ allele. Unfortunately, an initial attempt to verify an alteration in the 723SJ allele was unsuccessful.

Molecular and biochemical evidence presented in this work support the assertion that the *psrt* animals have an alteration in the *Prmt3* transcript. This evidence, together with the nature of the phenotype, leads to the conclusion that the *Prmt3* gene is responsible for the *psrt* phenotype.

CHAPTER X

FUTURE WORK

The findings from this work suggest that *Prmt3* is the gene responsible for the *psrt* phenotype. They do not, however, prove this assertion. In order to prove unequivocally that *Prmt3* is the gene responsible for the *psrt* phenotype, there are several experiments that need to be done. First of all, the 5' end of the 723SJ *Prmt3* transcript will have to be sequenced. Since RT-PCR and RACE failed to accomplish this, it will be necessary to sequence from genomic DNA. Even if another Southern blot (using different restriction enzymes) shows a rearrangement or deletion, sequence data will be necessary to define the nature of the mutation. It is also important to find the mutations that the TGCE heteroduplex analysis indicates are present in the 3' ends of the mutant transcripts. Further sequence analysis, utilizing either genomic DNA or cDNAs, will be needed. The larger transcripts detected in skeletal muscle and heart also must be characterized.

In order to prove that the biological function of the *Prmt3* gene is to methylate MBP, an additional Me-R assay must be performed, this time using both mutant strains. The experiment should be repeated using more brains (and, therefore, more MBP) so that the level of Me-R detected in the control animals is substantially above the detection limits for the procedure. And finally, either a gene knock-out from a wild type animal or a transgenic rescue of a mutant will serve as definitive proof that *Prmt3* is the gene responsible for the *psrt* phenotype in the mouse.

REFERENCES

Abraham JA, Mergia A, Whang JL, Tumolo A, Friedman J, Hjerrild KA, Gospodarowicz D, Fiddes JC (1986) Nucleotide sequence of a bovine clone encoding the angiogenic protein, basic fibroblast growth factor. Science 233(4763), 545-548

Abramovich C, Yakobson B, Chebath J, Revel M (1997) A protein-arginine methyltransferase binds to the intracytoplasmic domain of the IFNAR1 chain in the type I interferon receptor. EMBO J 16(2), 260-266

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1996) *Molecular Biology of the Cell*, 3rd Edition. (New York: Garland Publishing)

Aletta JM, Cimato TR, Ettinger MJ (1998) Protein methylation: a signal event in post translational modification. Trends Biochem Sci 23(3), 89-91

Alonso S, Minty A, Bourlet Y, Buckingham M (1986) Comparison of three actin coding sequences in the mouse; evolutionary relationships between the actin genes of warm-blooded vertebrates. J Mol Evol 23(1),11-22

Amur SG, Shanker G, Pieringer RA (1984) Regulation of myelin basic protein (arginine) methyltransferase by thyroid hormone in myelinogenic cultures of cells dissociated from embryonic mouse brain. J Neurochem 43(2), 494-498

Amur SG, Shanker G, Cochran JM, Ved HS, Pieringer RA (1986) Correlation between inhibition of myelin basic protein (arginine) methyltransferase by sinefungin and lack of compact myelin formation in cultures of cerebral cells from embryonic mice. J Neurosci Res 16, 367-376

Aris JP, Blobel G (1991) cDNA cloning and sequencing of human fibrillarin, a conserved nucleolar protein recognized by autoimmune antisera. Proc Natl Acad Sci USA 88, 931-935

Aubrey KR, Mitrovic AD, Vandenberg RJ (2000) Molecular basis for proton regulation of glycine transport by glycine transporter subtype 1b. Mol Pharmacol 58, 129-135

Baker ME (1999) TIP30, a cofactor for HIV-1 Tat-activated transcription, is homologous to short-chain dehydrogenases/reductases. Curr Biol 9(13), R471

Baker ME, Yan L, Pear MR (2000) Three-dimensional model of human TIP30, a coactivator for HIV-1 Tat-activated transcription, and CC3, a protein associated with metastasis suppression. Cell Mol Life Sci 57, 851-858

Baldwin GS, Carnegie PR (1971a) Isolation and partial characterization of methylated arginines from the encephalitogenic basic protein of myelin. Biochem J 123, 69-74

Baldwin GS, Carnegie PR (1971b) Specific enzymic methylation of an arginine in the experimental allergic encelphalomyelitis protein from human myelin. Science 171, 579-581

Ballabio A (1993) The rise and fall of positional cloning? Nat Genet 3, 277-279

Barbarese E, Carson JH, Braun PE (1978) Accumulation of the four myelin basic proteins in mouse brain during development. J Neurochem 31(4), 779-782

Barbarese E, Nielson ML, Carson JH (1983) The effect of the shiverer mutation on myelin basic protein expression in homozygous and heterozygous mouse brain. J Neurochem 40(6), 1680-1686

Bedford MT, Frankel A, Yaffe MB, Clarke S, Leder P, Richard S (2000) Arginine methylation inhibits the binding of proline-rich ligands to Src Hhomology 3, but not WW, domains. J Biol Chem 275(21), 16030-16036

Bell JA, Rinchik EM, Raymond S, Suffolk R, Jackson IJ (1995) A high-resolution map of the brown (*b, Tyrp 1)* deletion complex of mouse Chromosome 4. Mamm Genome 6, 389-395

Bergstrom RA, You Y, Erway LC, Lyon MF, Schimenti JC (1998) Deletion mapping of the head tilt *(het)* gene in mice: a vestibular mutation causing specific absence of otoliths. Genetics 150, 815-822

Biamonti G, Buvoli M, Bassi MT, Morandi C, Cobianchi F, Riva S (1989) Isolation of an active gene encoding human hnRNP protein A1. Evidence for alternative splicing. J Mol Biol 207(3), 491-503

Bird TD, Farrell DF, Sumi SM (1978) Brain lipid composition of the shiverer mouse: (genetic defect in myelin development). J Neurochem 31, 387-391

Boehm T (1998) Positional cloning and gene identification. Methods Companion Methods Enzymol 14, 152-158

Boggs JM, Rangaraj G, Koshy KM (1999) Analysis of the membrane-interacting domains of myelin basic protein by hydrophobic photolabeling. Biochim Biophys Acta 1417(2), 254-266

Bourre JM, Jacque C, Delassalle A, Nguyen-Legros J, Dumont O, Lachapelle F, Raoul M, Alvarez C, Baumann N (1980) Density profile and basic protein measurements in the myelin range of particulate material from normal developing mouse brain and from neurological mutants (jimpy; quaking; trembler; shiverer and its mld allele) obtained by zonal centrifugation. J Neurochem 35(2), 458-464

Bousquet-Antonelli C, Vanrobays E, Gélugne J-P, Caizergues-Ferrer M, Henry Y (2000) Rrp8p is a yeast nucleolar protein functionally linked to Gar1p and involved in pre-rRNA cleavage at site A2. RNA 6, 826-843

Bouvet P, Diaz J-J, Kindbeiter K, Madjar J-J, Amalric F (1998) Nucleolin interacts with several ribosomal proteins through its RGG domain. J Biol Chem 273(30), 19025-19029

Brahms H, Raymackers J, Union A, de Keyser F, Meheus L, Luhrmann R (2000) The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies. J Biol Chem 275(22), 17122-17129

Brilliant MH (1992) The mouse *pink-eyed dilution* locus: a model for aspects of Prader-Willi syndrome, Angelman syndrome, and a form of hypomelanosis of Ito. Mamm Genome 3, 187-191

Brilliant MH, Ching A, Nakatsu Y, Eicher EM (1994) The original *pink-eyed dilution* mutation (*p*) arose in Asiatic mice: implications for the H4 minor

histocompatibility antigen, *Myod1* regulation and the origin of inbred strains. Genetics 138(1), 203-211

Brostoff S, Eylar EH (1971) Localization of methylated arginine in the A1 protein from myelin. Proc Nat Acad Sci USA 68(4), 765-769

Buckler AJ, Chang DD, Graw SL, Brook JD, Haber DA, Sharp PA, Housman DE (1991) Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc Natl Acad Sci USA 88, 4005-4009

Burgess WH, Bizik J, Mehlman T, Quarto N, Rifkin DB (1991) Direct evidence for methylation of arginine residues in high molecular weight forms of basic fibroblast growth factor. Cell Regul 2(2), 87-93

Bussiere, DE, Muchmore SW, Dealwis CG, Schluckebier G, Nienaber VL, Edalji RP, Walter KA, Ladror US, Holzman TF, Abad-Zapatero C (1998) Crystal structure of ErmC', an rRNA methyltransferase which mediates antibiotic resistance in bacteria. Biochemistry 37, 7103-7112

Buvoli M, Biamonti G, Tsoulfas P, Bassi MT, Ghetti A, Riva S, Morandi C (1988) cDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms. Nucleic Acids Res 16(9), 3751-70

Calado A, Tomé FMS, Brais B, Rouleau GA, Kühn U, Wahle E, Carmo-Fonseca M (2000) Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. Hum Mol Genet 9(15), 2321-2328

Campagnoni CW, Carey GD, Campagnoni AT (1978) Synthesis of myelin basic proteins in the developing mouse brain. Arch Biochem Biophys 190(1), 118-125

Campagnoni AT, Pribyl TM, Campagnoni CW, Kampf K, Amur-Umarjee S, Landry CF, Handley VW, Newman SL, Garbay B, Kitamura K (1993) Structure and developmental regulation of Golli-mbp, a 105-kilobase gene that encompasses the myelin basic protein gene and is expressed in cells in the oligodendrocyte lineage in the brain. J Biol Chem 268(7), 4930-4938

Cao L, Goodin R, Wood D, Moscarello MA, Whitaker JN (1999) Rapid release and unusual stability of immunodominant peptide 45-89 from citrullinated myelin basic protein. Biochemistry 38, 6157-6163

Carson JH, Nielson ML, Barbarese E (1983) Developmental regulation of myelin basic protein expression in mouse brain. Dev Biol 96(2), 485-492

Casellas P, Jeanteur P (1978) Protein methylation in animal cells. I. Purification and properties of S-adenosyl-L-methionine:protein (arginine) N-methyltransferase from Krebs II ascites cells. Biochim Biophys Acta 519(1), 243-54

Chanderkar LP, Paik WK, Kim S (1986) Studies on myelin-basic-protein methylation during mouse brain development. Biochem J 240(2), 471-479

Chen D, Ma H, Hong H, Koh SS, Huang S-M, Schurter BT, Aswad DW, Stallcup MR (1999) Regulation of transcription by a protein methyltransferase. Science 284(5423), 2174-2177

Chen D, Huang S-M, Stallcup MR (2000) Synergistic, p160 coactivator-dependent enhancement of estrogen receptor function by CARM1 and p300. J Biol Chem 275(52), 40810-40816

Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18(24), 5294-5299

Church DM, Stotler CJ, Rutter JL, Murrell JR, Trofatter JA, Buckler AJ (1994) Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. Nat Genet 6, 98-105

Claverie J-M (1994) A streamlined random sequencing strategy for finding coding exons. Genomics 23, 575-581

Collins F (1992) Positional cloning: let's not call it reverse anymore. Nat Genet 1, 3-6

Cox RD, Hugill A, Shedlovsky A, Noveroske JK, Best S, Justice MJ, Lehrach H, Dove WF (1999) Contrasting effects of ENU induced embryonic lethal mutations of the quaking gene. Genomics 57(3), 333-341

Crang AJ, Jacobson W (1982) The relationship of myelin basic protein (arginine) methyltransferase to myelination in mouse spinal cord. J Neurochem 39(1), 244-247

Cujec TP, Okamoto H, Fujinaga K, Meyer J, Chamberlin H, Morgan DO, Peterlin BM (1997) The HIV transactivator TAT binds to the CDK-activating kinase and activates the phosphorylation of the carboxy-terminal domain of RNA polymerase II. Genes Dev 11, 2645-2657

Culiat CT, Stubbs L, Nicholls RD, Montgomery CS, Russell LB, Johnson DK, Rinchik EM (1993) Concordance between isolated cleft palate in mice and alterations within a region including the gene encoding the β_3 subunit of the type A γ-aminobutyric acid receptor. Proc Natl Acad Sci USA 90, 5105-5109

Culiat CT, Stubbs LJ, Montgomery CS, Russell LB, Rinchik EM (1994) Phenotypic consequences of deletion of the γ_3, α_5, or β_3 subunit of the type A γ-aminobutyric acid receptor in mice. Proc Natl Acad Sci USA 91, 2815-2818

Davis AP, Woychik RP, Justice MJ (1999) Effective chemical mutagenesis in FVB/N mice requires low doses of ethylnitrosourea. Mamm Genome 10, 308-310

DeBry RW, Seldin MF (1996) Human/mouse homology relationships. Genomics 33, 337-351

de Ferra F, Engh H, Hudson L, Kamholz J, Puckett C, Molineaux S, Lazzarini RA (1985) Alternative splicing accounts for the four forms of myelin basic protein. Cell 43(3 Pt 2), 721-727

Deibler GE, Martenson RE (1973) Determination of methylated basic amino acids with the amino acid analyzer. J Biol Chem 248(7), 2387-2391

Dhar MS, Johnson DK (1997) A microsatellite map of the *pink-eyed dilution* (*p*) deletion complex in mouse Chromosome 7. Mamm Genome 8, 143-145

Dhar MS, Webb LS, Smith L, Hauser L, Johnson D, West DB (2000) A novel ATPase on mouse chromosome 7 is a candidate gene for increased body fat. Physiol Genomics 4, 93-1000

Disa SG, Gupta A, Kim S, Paik WK (1986) Site specificity of histone H4 methylation by wheat germ protein-arginine N-methyltransferase. Biochemistry 25(9), 2443-8

Dono R, Texido G, Dussel R, Ehmke H, Zeller R (1998) Impaired cerebral cortex development and blood pressure regulation in FGF-2-deficient mice. EMBO J 17(15), 4213-4225

Duyk GM, Kim S, Myers RM, Cox DR (1990) Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. Proc Natl Acad Sci USA 87, 8995-8999

Evans J, Herdon H, Cairns W, O'Brien E, Chapman C, Terrett J, Gloger I (1999) Cloning, functional characterisation and population analysis of a variant form of the human glycine type 2 transporter. FEBS Lett 463, 301-306

Eylar EH, Thompson M (1969) Allergic encephalomyelitis: the physico-chemical properities of the basic protein encephalitogen from bovine spinal cord. Arch Biochem Biophys 129(2), 468-479

Farooqui JZ, Tuck M, Paik WK (1985) Purification and characterization of enzymes from Euglena gracilis that methylate methionine and arginine residues of cytochrome c. J Biol Chem 260(1), 537-45

Frankel A, Clarke S (1999) Rnase treatment of yeast and mammalian cell extracts affects *in vitro* substrate methylation by type I protein arginine *N*-methyltransferases. Biochem Biophys Res Comm 259, 391-400

Frankel A, Clarke S (2000) PRMT3 is a distinct member of the protein arginine N methyltransferase family. Conferral of substrate specificity by a zinc-finger domain. J Biol Chem 275(42), 32974-32982

Frost BF, Park KS, Tuck M, Disa S, Kim S, Paik WK (1989) Site-specificity of histone H1 methylation by two H1-specific protein-lysine *N*-methyltransferases from *Euglena gracilis*. Int J Biochem 21(9), 1061-70

Gao Q, Yeung ES (2000) High-throughput detection of unknown mutations by using multiplexed capillary electrophoresis with poly(vinylpyrrolidone) solution. Anal Chem 72(11), 2499-2506

Gardner JM, Nakatsu Y, Gondo Y, Lee S, Lyon MF, King RA, Brilliant MH (1992) The mouse pink-eyed dilution gene: association with human Prader-Willi and Angelman syndromes. Science 257(5073), 1121-1124

Gary JD, Lin W-J, Yang MC, Herschman HR, Clarke S (1996) The predominant protein-arginine methyltransferase-from *Saccharomyces cerevisiae.* J Biol Chem 271(21), 12585-12594

Gary JD, Clarke S (1998) RNA and protein interactions modulated by protein arginine methylation. Prog Nucleic Acid Res Mol Biol 61, 65-131

Geerlings A, López-Corcuera B, Aragón C (2000) Characterization of the interactions between the glycine transporters GLYT1 and GLYT2 and the SNARE protein syntaxin 1A. FEBS Lett 470, 51-54

Gegelashvili G, Schousboe A (1997) High affinity glutamate transporters: regulation of expression and activity. Mol Pharmacol 52, 6-15

Ghosh SK, Rawal N, Syed SK, Paik WK, Kim S (1991) Enzymic methylation of myelin basic protein in myelin. Biochem J 275, 381-387

Gluecksohn-Waelsch S (1979) Genetic control of morphogenetic and biochemical differentiation: lethal albino deletions in the mouse. Cell 16(2), 225-37

Gu H, Park SH, Park GH, Lim IK, Lee H-W, Paik WK, Kim S (1999) Identification of highly methylated arginine residues in an endogenous 20-kDa polypeptide in cancer cells. Life Sci 65(8), 737-745

Haas S, Steplewski A, Siracusa LD, Amini S, Khalili K (1995) Identification of a sequence-specific single-stranded DNA binding protein that suppresses transcription of the mouse myelin basic protein gene. J Biol Chem 270(21), 12503-12510

Hagmann M (1999) How chromatin changes its shape. Science 285(5431) 1200-1203

Haldane JBS, Sprunt AD, Haldane NM (1915) Reduplication in mice. J Genet 5, 133-135

Harrich D, Ulich C, García-Martínez LF, Gaynor R (1997) Tat is required for efficient HIV-1 reverse transcription. EMBO J 16(6), 1224-1235

Henry MF, Silver PA (1996) A novel methyltransferase (Hmt1p) modifies poly(A)$^+$ - RNA-binding proteins. Mol Cell Biol 16(7), 3668-3678

Herrmann CH, Rice AP (1995) Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. J Virol 69(3), 1612-1620

Hitotsumachi S, Carpenter DA, Russell WL (1985) Dose-repetition increases the mutagenic effectiveness of N-ethyl-N-nitrosourea in mouse spermatogonia. Proc Natl Acad Sci USA 82(19), 6619-21

Hoek KS, Kidd GJ, Carson JH, Smith R (1998) hnRNP A2 selectively binds the cytoplasmic transport sequence of myelin basic protein mRNA. Biochemistry 37(19), 7021-7029

Hogan EL, Greenfield S (1984) Animal models of genetic disorders of myelin. In *Myelin*, 2nd ed., P. Morell, ed. (New York: Plenum Press), pp 489-534

Holdener BC, Thomas JW, Schumacher A, Potter MD, Rinchik EM, Sharan SK, Magnuson T (1995) Physical localization of *eed:* a region of mouse Chromosome 7 required for gastrulation. Genomics 27, 447-456

Horiuchi M, El Far O, Betz H (2000) Ulip6, a novel unc-33 and dihydropyrimidinase related protein highly expressed in developing rat brain. FEBS Lett 480, 283-286

Huang J-D, Cope MJTV, Mermall V, Strobel MC, Kendrick-Jones J, Russell LB, Mooseker MS, Copeland NG, Jenkins NA (1998a) Molecular genetic dissection of mouse unconventional myosin-VA: head region mutations. Genetics 148, 1951-1961

Huang J-D, Mermall V, Strobel MC, Russell LB, Mooseker MS, Copeland NG, Jenkins NA (1998b) Molecular genetic dissection of mouse unconventional myosin-VA: tail region mutations. Genetics 148, 1963-1972

Hyun Y-L, Lew DB, Park SH, Kim C-W, Paik WK, Kim S (2000) Enzymic methylation of arginyl residues in -Gly-Arg-Gly peptides. Biochem J 348, 573-578

Jacque C, Delassalle A, Raoul M, Baumann N (1983) Myelin basic protein deposition in the optic and sciatic nerves of dysmyelinating mutants quaking, jimpy, trembler, MLD, and shiverer during development. J Neurochem 41 (5), 1335-1340

Jeang K-T, Xiao H, Rich EA (1999) Multifaceted activities of the HIV-1 transactivator of transcription, Tat. J Biol Chem 274(41), 28837-28840

Jensen L, Kuff EL, Wilson SH, Steinberg AD, Klinman DM (1988) Antibodies from patients and mice with autoimmune diseases react with recombinant hnRNP core protein A1. J Autoimmun 1(1), 73-83

Ji W, Herron B, Jones JM, Jenkins NA, Gilbert DJ, Copeland NG, Swank R, Flaherty L, Meisler MH (1999) Identification of genes within the *Krd* deletion on mouse Chromosome 19. Mamm Genome 10, 399-401

Johnson DK, Hand RE Jr., Rinchik EM (1989) Molecular mapping within the mouse albino-deletion complex. Proc Natl Acad Sci USA 86, 8862-8866

Johnson DK, Stubbs LJ, Culiat CT, Montgomery CS, Russell LB, Rinchik EM (1995) Molecular analysis of 36 mutations at the mouse *pink-eyed dilution* (*p*) locus. Genetics 141, 1563-1571

Jonas P, Bischofberger J, Sandkùhler J (1998) Corelease of two fast neurotransmitters at a central synapse. Science 281, 419-424

Justice MJ, Noveroske JK, Weber JS, Zheng B, Bradley A (1999) Mouse ENU mutagenesis. Hum Mol Genet 8(10), 1955-1963

Justice MJ, Carpenter DA, Favor J, Neuhauser-Klaus A, Hrabe de Angelis M, Soewarto D, Moser A, Cordes S, Miller D, Chapman V, Weber JS, Rinchik EM, Hunsicker PR, Russell WL, Bode VC (2000) Effects of ENU dosage on mouse strains. Mamm Genome 11, 484-488

Kamholz J, Spielman R, Gogolin K, Modi W, O'Brien S, Lazzarini R (1987) The human myelin-basic-protein gene: chromosomal localization and RFLP analysis. Am J Hum Genet 40(4), 365-373

Katsanis N, Yaspo M-L, Fisher EMC (1997) Identification and mapping of a novel human gene, HRMT1L1, homologous to the rat protein arginine N-methyltransferase 1 (PRMT1) gene. Mamm Genome 8, 526-529

Keen NJ, Churcher MJ, Karn J (1997) Transfer of Tat and release of TAR RNA during the activation of the human immunodeficiency virus type-1 transcription elongation complex. EMBO J 16(17), 5260-5272

Kies MW, Martenson RE, Deibler GE (1972) Myelin basic proteins. Adv Exp Med Biol 32, 201-214

Kim S, Chanderkar LP, Ghosh SK, Park J-O, Paik WK (1988) Enzymatic methylation of arginine residue in myelin basic protein. Adv Exp Med Biol 231, 327-340

Kim U-J, Birren BW, Slepak T, Mancino V, Boysen C, Kang H-L, Simon MI, Shizuya H (1996) Construction and characterization of a human bacterial artificial chromosome library. Genomics 34, 213-218

Kim J, Carver EA, Stubbs L (1997a) Amplification and sequencing of end fragments from bacterial artificial chromosome clones by single-primer polymerase chain reaction. Anal Biochem 253(2), 272-275

Kim S, Lim IK, Park G-H, Paik WK (1997b) Biological methylation of myelin basic protein: enzymology and biological significance. Int J Biochem Cell Biol 29(5), 743-751

Kim S, Merrill BM, Rajpurohit R, Kumar A, Stone KL, Papov VV, Schneiders JM, Szer W, Wilson SH, Paik WK, Williams KR (1997c) Identification of N^G-methylarginine residues in human heterogeneous RNP protein A1: Phe/Gly-Gly-Gly-Arg-Gly-Gly-Gly/Phe is a preferred recognition motif. Biochemistry 36, 5185-5192

Kim JB, Yamaguchi Y, Wada T, Handa H, Sharp PA (1999) Tat-SF1 protein associates with RAP30 and human SPT5 proteins. Molec Cell Biol 19(9), 5960-5968

Kirschner DA, Ganser AL (1980) Compact myelin exists in the absence of basic protein in the shiverer mutant mouse. Nature 283, 207-210

Klein S, Carroll JA, Chen Y, Henry MF, Henry PA, Ortonowshi IE, Pintucci G, Beavis RC, Burgess WH, Rifkin DB (2000) Biochemical analysis of the arginine methylation of high molecular weight fibroblast growth factor-2. J Biol Chem 275(5), 3150-7

Koh SS, Chen D, Lee YH, Stallcup MR (2001) Synergistic enhancement of nuclear receptor function by p160 coactivators and two coactivators with protein methyltransferase activities. J Biol Chem 276(2), 1089-1098

Kolehmainen E, Sormunen R (1998) Myelin basic protein induces morphological changes in the endocrine pancreas. Pancreas 16(2) 176-188

Kolehmainen E, Knip M, Leppaluoto J (1990) Myelin basic protein stimulates insulin and glucagon secretion from rat pancreatic islets in vitro and in vivo. Acta Physiol Scand 139(3), 493-501

Krizman DB, Berget SM (1993) Efficient selection of 3'-terminal exons from vertebrate DNA. Nucleic Acids Res 21(22), 5198-5202

Kruger GM, Diemel LT, Copelman CA, Cuzner ML (1999) Myelin basic protein isoforms in myelinating and remyelinating rat brain aggregate cultures. J Neurosci Res 56(3), 241-247

Kurokawa T, Sasada R, Iwane M, Igarashi K (1987) Cloning and expression of cDNA encoding human basic fibroblast growth factor. FEBS Lett 213(1), 189-94

Labahn J, Granzin J, Schluckebier G, Robinson DP, Jack WE, Schildkraut I, Saenger W (1994) Three-dimensional structure of the adenine-specific DNA methyltransferase M *Taq* I in complex with the cofactor *S*-adenosylmethionine. Proc Natl Acad Sci USA 91, 10957-10961

Lee HW, Kim S, Paik WK (1977) *S*-Adenosylmethionine: protein-arginine methyltransferase. Purification and mechanism of the enzyme. Biochemistry 16(1), 78-85

Lee S-T, Nicholls RD, Schnur RE, Guida LC, Lu-Kuo J, Spinner NB, Zackai EH, Spritz RA (1994) Diverse mutations of the *P* gene among African-Americans with

type II (tyrosinase-positive) oculocutaneous albinism (OCA2). Hum Mol Genet 3(11), 2047-2051

Lee S-T, Nicholls RD, Jong MTC, Fukai K, Spritz RA (1995) Organization and sequence of the human *P* gene and identification of a new family of transport proteins. Genomics 26, 354-363

Lee JH, Cook JR, Pollack BP, Kinzy TG, Norris D, Pestka S (2000) Hsl7p, the yeast homologue of human JBP1, is a protein methyltransferase. Biochem Biophys Res Commun 274(1), 105-11

Lees MB, Brostoff SW (1984) Proteins of Myelin. In Myelin, 2nd ed., P Morell, ed. (New York: Plenum Press), pp 197-224

Lesch KP, Heils A, Riederer P (1996) The role of neurotransporters in excitotoxicity, neuronal cell death, and other neurodegenerative processes. J Mol Med 74(7), 365-78

Lin WJ, Gary JD, Yang MC, Clarke S, Herschman HR (1996) The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. J Biol Chem 271(25), 15034-15044

Lindsay EA, Botta A, Jurecic V, Carattini-Rivera S, Cheah YC, Rosenblatt HM, Bradley A, Baldini A (1999) Congenital heart disease in mice deficient for the DiGeorge syndrome region. Nature 401(6751), 379-83

Liu Q, Dreyfuss G (1995) *In vivo* and *in vitro* arginine methylation of RNA-binding proteins. Mol Cell Biol 15(5), 2800-2808

Liu Y, Edwards RH (1997) The role of vesicular transport proteins in synaptic transmission and neural degeneration. Annu Rev Neurosci 20, 125-56

Liu Y, Thor A, Shtivelman E, Cao Y, Tu G, Heath TD, Debs RJ (1999) Systemic gene delivery expands the repertoire of effective antiangiogenic agents. J Biol Chem 274(19) 13338-13344

Lovett M, Kere J, Hinton LM (1991) Direct selection: a method for the isolation of cDNAs encoded by large genomic regions. Proc Natl Acad Sci USA 88, 9628-9632

Lyon MF, Morris T (1966) Mutation rates at a new set of specific loci in the mouse. Genet Res 7, 12-17

Lyon MF, King TR, Gondo Y, Gardner JM, Nakatsu Y, Eicher EM, Brilliant MH (1992) Genetic and molecular analysis of recessive alleles at the *pink-eyed dilution* (*p*) locus of the mouse. Proc Natl Acad Sci USA 89, 6968-6972

Maatta JA, Coffey ET, Hermonen JA, Salmi AA, Hinkkanen AE (1997) Detection of myelin basic protein isoforms by organic concentration. Biochem Biophys Res Commun 238(2), 498-502

Mackay JP, Crossley M (1998) Zinc fingers are sticking together. TIBS J 23, 1-4

Marker PC, Seung K, Bland AE, Russell LB, Kinsley DM (1997) Spectrum of *Bmp5* mutations from germline mutagenesis experiments in mice. Genetics 145(2), 435-443

Martinez-Maza R, Poyatos I, Lopez-Corcuera B, Nunez E, Gimenez C, Zafra F, Aragon C (2001) The role of *N*-glycosylation in transport to the plasma membrane and

sorting of the neuronal glycine transporter GLYT2. J Biol Chem 276(3) 2168-2173

Mathisen PM, Pease S, Garvey J, Hood L, Readhead C (1993) Identification of an embryonic isoform of myelin basic protein that is expressed widely in the mouse Embryo. Proc Natl Acad Sci USA 90(21), 10125-10129

Mears WE, Lam V, Rice SA (1995) Identification of nuclear and nucleolar localization signals in the herpes simplex visur regulatory protein ICP27. J Virol 69(2), 935-947

Mears WE, Rice SA (1996) The RGG box motif of the herpes simplex virus ICP27 protein mediates an RNA-binding activity and determines in vivo methylation. J Virol 70(11), 7445-7453

Metallinos DL, Oppenheimer AJ, Rinchik EM, Russell LB, Dietrich W, Tilghman SM (1994) Fine structure mapping and deletion analysis of the murine *piebald* locus. Genetics 136, 217-223

Michael WM, Choi M, Dreyfuss G (1995) A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. Cell 83(3), 415-22

Minota S, Jarjour WN, Suzuki N, Nojima Y, Roubey RA, Mimura T, Yamada A, Hosoya T, Takaku F, Winfield JB (1991) Autoantibodies to nucleolin in systemic lupus erythematosus and other diseases. J Immunol 146(7), 2249-52

Montero A, Okada Y, Tomita M, Ito M, Tsurukami H, Nakamura T, Doetschman T, Coffin JD, Hurley MM (2000) Disruption of the *fibroblast growth factor-2* gene

results in decreased bone mass and bone formation. J Clin Invest 105(8), 1085-93

Morrow JA, Collie IT, Dunbar DR, Walker GB, Shahid M, Hill DR (1998) Molecular cloning and functional expression of the human glycine transporter *GlyT2* and chromosomal lcalisation of the gene in the human genome. FEBS Lett 439, 334-340

Moscarello MA, Wood DD, Ackerley C, Boulias C (1994) Myelin in multiple sclerosis is developmentally immature. J Clin Invest 94, 146-154

Mueller PR, Wold B (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. Science 246, 780-786

Muller D, Rehbein M, Baumeister H, Richter D (1992) Molecular characterization of a novel rat protein structurally related to poly(A) binding proteins and the 70K protein of the U1 small nuclear ribonucleoprotein particle. Nucleic Acids Res 20(7), 1471-5

Najbauer J, Johnson BA, Aswad DW (1992) Analysis of stable protein methylation in cultured cells. Arch Biochem Biophys 293(1), 85-92.

Najbauer J, Johnson BA, Young AL, Aswad DW (1993) Peptides with sequences similar to glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins. J Biol Chem 268(14), 10501-10509

Nemeth A, Krause S, Blank D, Jenny A, Jeno P, Lustig A, Wahle E (1995) Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II. Nucleic Acids Res 23(20), 4034-41

Nicholls RD, Gottlieb, Russell LB, Davda M, Horsthemke B, Rinchik EM (1993) Evaluation of potential models for imprinted and nonimprinted components of human chromosome 15q11-q13 syndromes by fine-structure homology mapping in the mouse. Proc Natl Acad Sci USA 90, 2050-2054

Nichols RC, Wang XW, Tang J, Hamilton BJ, High FA, Herschman HR, Rigby WF (2000) The RGG domain in hnRNP A2 affects subcellular localization. Exp Cell Res 256(2) 522-532

Nicoll RA, Malenka RC (1998) A tale of two transmitters. Science 281(5375), 360-361

Nicoloso M, Caizergues-Ferrer M, Michot B, Azum MC, Bachellerie JP (1994) U20, a novel small nucleolar RNA, is encoded in an intron of the nucleolin gene in mammals. Mol Cell Biol 14(9), 5766-76

Niewmierzycka A, Clarke S (1999) *S*-Adenosylmethionine-dependent methylation in *Saccharomyces cerevisiae.* J. Biol Chem 274(2), 814-824

Niswander L, Yee D, Rinchik EM, Russell LB, Magnuson T (1988) The albino deletion complex and early postimplantation survival in the mouse. Development 102(1), 45-53

Noveroske JK, Weber JS, Justice MJ (2000) The mutagenic action of *N*-ethyl-*N*-nitrosourea in the mouse. Mamm Genome 11, 478-483

O'Brien TP, Metallinos DL, Chen H, Shin MK, Tilghman SM (1996) Complementation mapping of skeletal and central nervous system abnormalities in mice of the *piebald* deletion complex. Genetics 143, 447-461

Ochs RL, Lischwe MA, Spohn WH, Busch H (1985) Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. Biol Cell 54(2), 123-33

Omlin FX, Webster HD, Palkovits CG, Cohen SR (1982) Immunocytochemical localization of basic protein in major dense line regions of central and peripheral myelin. J Cell Biol 95(1), 242-248

Ortega S, Ittmann M, Tsang SH, Ehrlich M, Basilico C (1998) Neuronal defects and delayed wound healing in mice lacking *fibroblast growth factor 2*. Proc Natl Acad Sci USA 95, 5672-5677

Parada CA, Roeder RG (1996) Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. Nature 384, 375-378

Parimoo S, Patanjali SR, Shukla H, Chaplin DD, Weissman SM (1991) cDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. Proc Natl Acad Sci USA 88, 9623-9627

Parimoo S, Patanjali SR, Kolluri R, Xu H, Wei H, Weissman SM (1995) cDNA selection and other approaches in positional cloning. Anal Biochem 228, 1-17

Park J, Greenstein JI, Paik WK, Kim S (1989) Studies on protein methyltransferase in human cerebrospinal fluid. J Mol Neurosci 1, 151-157

Pawlak MR, Scherer CA, Chen J, Roshon MJ, Ruley HE (2000) Arginine *N* methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable. Mol Cell Biol 20(13), 4859-4869

Pennisi E (1997) Opening the way to gene activity. Science 275(5297), 155-157

Plotnikov AN, Hubbard SR, Schlessinger J, Mohammadi M (2000) Crystal structures of two FGF-FGFR complexes reveal the determinants of ligand-receptor specificity. Cell 101(4), 413-24

Pollack BP, Kotenko SV, He W, Izotova LS, Barnoski BL, Pestka S (1999) The human homologue of the yeast proteins Skb1 and Hs17p interacts with Jak kinases and contains protein methyltransferase activity. J Biol Chem 274(44), 31531-31542

Pombo PM, Barettino D, Ibarrola N, Vega S, Rodriguez-Pena A (1999) Stimulation of the myelin basic protein gene expression by 9-cis-retinoic acid and thyroid hormone: activation in the context of its native promoter. Brain Res Mol Brain Res 64(1), 92-100

Ponce J, Biton B, Benavides J, Avenet P, Aragon (2000) Transmembrane domain III plays an important role in ion binding and permeation in the glycine transporter GLYT2. J Biol Chem 275(18), 13856-13862

Popko B, Puckett C, Lai E, Shine HD, Readhead C, Takahashi N, Hunt SW III, Sidman RL, Hood L (1987) Myelin deficient mice: expression of myelin basic protein and generation of mice with varying levels of myelin. Cell 48, 713-721

Pothos EN, Larsen KE, Krantz DE, Liu Y-J, Haycock JW, Setlik W, Gershon MD, Edwards RH, Sulzer D (2000) Synaptic vesicle transporter expression regulates vesicle phenotype and quantal size. J Neurosci 20(19), 7297-7306

Potter MD, Rinchik EM (1993) Deletion mapping of the chocolate *(cht)* locus within the *Fes-Hbb* region of mouse Chromosome 7. Mam Genome 4, 46-48

Potter MD, Klebig ML, Carpenter DA, Rinchik EM (1995) Genetic and physical mapping of the *fitness I (fitI)* locus within the *Fes-Hbb* region of mouse Chromosome 7. Mamm Genome 6, 70-75

Prasad K, Barouch W, Martin BM, Greene LE, Eisenberg E (1995) Purification of a new clathrin assembly protein from bovine brain coated vesicles and its identification as myelin basic protein. J Biol Chem 270(51), 30551-30556

Pribyl TM, Campagnoni CW, Kampf K, Kashima T, Handley VW, McMahon J, Campagnoni AT (1993) The human myelin basic protein gene is included within a 179-kilobase transcription unit: expression in the immune and central nervous systems. Proc Natl Acad Sci USA 90(22), 10695-10699

Raine CS (1984a) Morphology of myelin and myelination. In Myelin, 2nd ed., P. Morell, ed. (New York: Plenum Press), pp 1-50

Raine CS (1984b) The neuropathology of myelin diseases. In Myelin, 2nd ed., P. Morell, ed. (New York: Plenum Press), pp 259-310

Rajpurohit R, Paik WK, Kim S (1994a) Effect of enzymic methylation of heterogeneous ribonucleoprotein particle A1 on its nucleic-acid binding and controlled proteolysis. Biochem J 304, 903-909

Rajpurohit R, Lee SO, Park JO, Paik WK, Kim S (1994b) Enzymatic methylation of recombinant heterogeneous nuclear RNP protein A1. J Biol Chem 269(2), 1075-1082

Ramirez-Solis R, Liu P, Bradley A (1995) Chromosome engineering in mice. Nature 378(6558), 720-4

Ramsay M, Colman M-A, Stevens G, Zwane E, Kromberg J, Farrall M, Jenkins T (1992) The tyrosinase-positive oculocutaneous albinism locus maps to chromosome 15q11.2-q12. Am J Hum Genet 51, 879-884

Rawal N, Paik WK, Kim S (1991) An enzyme-linked immunosorbent assay for myelin basic protein-specific protein methylase I. J Neurosci Methods 37, 133-140

Rawal N, Rajpurohit R, Lischwe MA, Williams KR, Paik WK, Kim S (1995) Structural specificity of substgrate for *S*-adenosylmethionine: protein arginine *N*-methyltransferases. Biochim Biophys Acta 1248, 11-18

Rawal N, Lee YJ, Whitaker JN, Park, JO, Paik WK, Kim S (1995b) Urinary excretion of NG-dimethylarginines in multiple sclerosis patients: preliminary observations. J Neurol Sci 129(2), 186-191

Readhead C, Popko B, Takahashi N, Shine HD, Saavedra RA, Sidman RL, Hood L (1987) Expression of a myelin basic protein gene in transgenic shiverer mice: correction of the dysmyelinating phenotype. Cell 48, 703-712

Rekling JC, Funk GD, Bayliss DA, Dong X-W, Feldman JL (2000) Synaptic control of motoneuronal excitability. Physiol Rev 80(2), 767-852

Rho J, Choi S., Seong YR, Cho WK, Kim SH, Im DS (2001) *Prmt5*, which forms distinct homo-oligomers, is a member of the protein-arginine methyltransferase family. J Biol Chem 276(14), 11393-11401

Rikke BA, Johnson DK, Johnson TE (1997) Murine *albino*-deletion complex: high-resolution microsatellite map and genetically anchored YAC framework map. Genetics 147, 787-799

Riley J, Butler R, Ogilvie D, Finniear R, Jenner D, Powell S, Anand R, Smith JC, Markham AF (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res 18(10), 2887-2890

Rinchik EM (1991) Chemical mutagenesis and fine-structure functional analysis of the mouse genome. TIG 7(1), 15-21

Rinchik EM (1994) Molecular genetics of the *brown (b)*-locus region of mouse chromosome 4. II. Complementation analyses of lethal *brown* deletions. Genetics 137, 855-865

Rinchik EM (2000) Developing genetic reagents to facilitate recovery, analysis, and maintenance of mouse mutations. Mamm Genome 11, 489-499

Rinchik EM, Carpenter DA (1993) *N*-Ethyl-*N*-nitrosourea-induced postimplantation-lethal mutations within the *pid-Hbb* region of mouse chromosome 7. Mamm Genome 4(7), 349-353

Rinchik EM, Carpenter DA (1999) *N*-Ethyl-*N*-nitrosourea mutagenesis of a 6- to 11-cM subregion of the *Fah-Hbb* interval of mouse chromosome 7: completed

testing of 4557 gametes and deletion mapping and complementation analysis of 31 mutations. Genetics 152, 373-383

Rinchik EM, Russell LB (1990) Germ-line deletion mutations in the mouse: tools for intensive functional and physical mapping of regions of the mammalian genome. In Genetic and physical Mapping. Volume I: Genome Analysis, K.E. Davies and S.M.Tilghman, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp 121-158

Rinchik EM, Machanoff R, Cummings CC, Johnson DK (1989) Molecular cloning and mapping of the ecotropic leukemia provirus *Emv-23* provides molecular access to the albino-deletion complex in mouse chromosome 7. Genomics 4, 251-258

Rinchik EM, Carpenter DA, Selby PB (1990) A strategy for fine-structure functional analysis of a 6- to 11-centimorgan region of mouse chromosome 7 by high-efficiency mutagenesis. Proc Natl Acad Sci USA 87, 896-900

Rinchik EM, Bultman SJ, Horsthemke B, Lee S-T, Strunk KM, Spritz RA, Avidano KM, Jong MTC, Nicholls RD (1993a) A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. Nature 361, 72-76

Rinchik EM, Carpenter DA, Long CL (1993b) Deletion mapping of four loci defined by *N*-Ethyl-*N*-Nitrosourea-induced postimplantation-lethal mutations within the *pid-Hbb* region of mouse chromosome 7. Genetics 135, 1117-1123

Rinchik EM, Bell JA, Hunsicker PR, Friedman JM, Jackson IJ, Russell LB (1994) Molecular genetics of the *brown (b)*-locus region of mouse chromosome 4. I.

Origin and molecular mapping of radiation- and chemical-induced lethal *brown* deletions. Genetics 137, 845-854

Rinchik EM, Carpenter DA, Handel MA (1995) Pleiotropy in microdeletion syndromes: Neurologic and spermatogenic abnormalities in mice homozygous for the p^{6H} deletion are likely due to dysfunction of a single gene. Proc Natl Acad Sci USA 92, 6394-6398

Roach A, Boylan K, Horvath S, Prusiner SB, Hood LE (1983) Characterization of cloned cDNA representing rat myelin basic protein: absence of expression in brain of shiverer mutant mice. Cell 34, 799-806

Rosemblat S, Durham-Pierre D, Gardner JM, Nakatsu Y, Brilliant MH, Orlow SJ (1994) Identification of a melanosomal membrane protein encoded by the *pink-eyed dilution* (type II oculocutaneous albinism) gene. Proc Natl Acad Sci USA 91, 12071-12075

Rosenbluth J (1980a) Central myelin in the mouse mutant shiverer. Comp Neurol (3), 639-648

Rosenbluth J (1980b) Peripheral myelin in the mouse mutant Shiverer. J Comp Neurol. 193(3), 729-739

Royer-Pokora B, Kunkel LM, Monaco AP, Goff SC, Newburger PE, Baehner RL, Cole FS, Curnutte JT, Orkin SH (1986) Cloning the gene for an inherited human disorder--chronic granulomatous disease--on the basis of its chromosomal location. Nature 322(6074), 32-8

Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, Grunstein M (1996) HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. Proc Natl Acad Sci USA 93. 14503-14508

Russell ES (1949) A quantitative histological study of the pigment found in the coat-color mutants of the house mouse. IV. The nature of the effects of genic substitution in five major allelic series. Genetics 34, 146-166

Russell WL (1951) X-ray induced mutations in mice. Cold Spring Harbor Symp Quant Biol 16, 327-336

Russell LB (1971) Definition of functional units in a small chromosomal segment of the mouse and its use in interpreting the nature of radiation-induced mutations. Mutation Res 11, 107-123

Russell LB (1979) Analysis of the *albino*-locus region of the mouse. II. Fractional mutants. Genetics 91, 141-147

Russell LB (1989) Functional and structural analyses of mouse genomic regions screened by the morphological specific-locus test. Mutat Res 212, 23-32

Russell LB (2001) Effects of male germ-cell stage on the frequency, nature, and spectrum of induced specific-locus mutations in the mouse. Genetica, in press

Russell LB, Raymer GC (1979) Analysis of the *albino*-locus region of the mouse. III. Time of death of prenatal lethals. Genetics 92, 205-213

Russell LB, Rinchik EM (1987) Genetic and molecular characterization of genomic regions surrounding specific loci of the mouse. In *Banbury Report 28:*

Mammalian Cell Mutagenesis (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp 109-121

Russell LB, Russell WL (1992) Frequency and nature of specific-locus mutations induced in female mice by radiations and chemicals: a review. Mutat Res 296, 107-127

Russell WL, Russell LB, Kelly EM (1958) Radiation dose rate and mutation frequency. Science 128(3338), 1546-50

Russell LB, Russell WL, Kelly EM (1979a) Analysis of the *albino*-locus region of the mouse. I. Origin and viability. Genetics 91, 127-139

Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC, Phipps EL (1979b) Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. Proc Natl Acad Sci USA 76(11), 5818-5819

Russell LB, Selby PB, von Halle E, Sheridan W, Valcovic L (1981) The mouse specific-locus test with agents other than radiations. Interpretation of data and recommendations for future work. Mutat Res 86, 329-354

Russell LB, Montgomery CS, Raymer GC (1982a) Analysis of the albino-locus region of the mouse: IV. Characterization of 34 deficiencies. Genetics 100, 427-453

Russell WL, Hunsicker PR, Raymer GD, Steele MH, Stelzner KF, Thompson HM (1982b) Dose-response curve for ethylnitrosourea-induced specific-locus mutations in mouse spermatogonia. Proc Natl Acad Sci USA 79(11), 3589-91

Russell WL, Hunsicker PR, Carpenter DA, Cornett CV, Guinn GM (1982c) Effect of dose fractionation on the ethylnitrosourea induction of specific-locus mutations in mouse spermatogonia. Proc Natl Acad Sci USA 79(11), 3592-3

Russell LB, Russell WL, Rinchik EM, Hunsicker PR (1990) Factors affecting the nature of induced mutations. In *Banbury Report 34: Biology of Mammalian Germ Cell Mutagenesis* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp 271-289

Russell LB, Montgomery CS, Cacheiro NLA, Johnson DK (1995) Complementation analyses for 45 mutations encompassing the *pink-eyed dilution* (*p*) locus of the mouse. Genetics 141, 1547-1562

Sambrook J, Fritsch EF, Maniatis, T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press)

Sandri-Goldin RM (1994) Properties of an HSV-1 regulatory protein that appears to impair host cell splicing. Infect Agents Dis 3(2-3), 59-67

Sandri-Goldin RM (1998) ICP27 mediates HSV RNA export by shuttling through a leucine-rich nuclear export signal and binding viral intronless RNAs through an RGG motif. Genes Dev 12, 868-879

Schellekens GA, de Jong BAW, van den Hoogen FHJ, van de Putte LBA, van Venrooij WJ (1998) Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 101(1), 273-281

Schimenti J, Bucan M (1998) Functional genomics in the mouse: phenotype-based mutagenesis screens. Genome Res 8, 698-710

Schimenti JC, Libby BJ, Bergstrom RA, Wilson LA, Naf D, Tarantino LM, Alavizadeh A, Lengeling A, Bucan M (2000) Interdigitated deletion complexes on mouse chromosome 5 induced by irradiation of embryonic stem cells. Genome Res 10, 1043-1050

Schmidt-Zachmann MS, Nigg EA (1993) Protein localization to the nucleolus: a search for targeting domains in nucleolin. J Cell Sci 105(3), 799-806

Scorilas A, Black MH, Talieri M, Diamandis EP (2000) Genomic organization, physical mapping, and expression analysis of the human *protein arginine methyltransferase 1* gene. Biochem Biophys Res Comm 278, 349-359

Scott, H. S.; Antonarakis, S. E.; Lalioti, M. D.; Rossier, C.; Silver, P. A.; Henry, M. F (1998) Identification and characterization of two putative human arginine methyltransferases (HRMT1L1 and HRMT1L2). Genomics 48, 330-340

Scrable HJ, Johnson DK, Rinchik EM, Cavenee WK (1990) Rhabdomyosarcoma-associated locus and *MYOD1* are syntenic but separate loci on the short arm of human chromosome 11. Proc Natl Acad Sci USA 87, 2182-2186

Shen EC, Henry MF, Weiss VH, Valentini SR, Silver PA, Lee MS (1998) Arginine methylation facilitates the nuclear export of hnRNP proteins. Genes Dev 12, 679-691

Shizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA

in *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci USA 89, 8794-8797

Shtivelman E (1997) A link between metastasis and resistance to apoptosis of variant small cell lung carcinoma. Oncogene J 14, 2167-2173

Sibilia M, Wagner EF (1995) Strain-dependent epithelial defects in mice lacking the EGF receptor. Science 269(5221), 234-8

Sidman RL, Conover CS, Carson JH (1985) Shiverer gene maps near the distal end of chromosome 18 in the house mouse. Cytogen Cell Genet J 39, 241-245

Siebel CW, Guthrie C (1996) The essential yeast RNA binding protein Np13p is methylated. Proc Natl Acad Sci USA 93, 13641-13646

Silver LM (1995) *Mouse Genetics*. (New York: Oxford University Press)

Silvers WK (1979) *The coat colors of mice*. (New York: Springer-Verlag)

Smith JJ, Rucknagel KP, Schierhorn A, Tang J, Nemeth A, Linder M, Herschman HR, Wahle E (1999) Unusual sites of arginine methylation in poly(A)-binding protein II and *in vitro* methylation by protein arginine methyltransferases PRMT1 and PRMT3. J Biol Chem 274(19), 13229-13234

Srivastava M, McBride OW, Fleming PJ, Pollard HB, Burns AL (1990) Genomic organization and chromosomal localization of the human *nucleolin* gene. J Biol Chem 265(25), 14922-31

Stallcup MR, Chen D, Koh SS, Ma H, Lee Y-H, Li H, Schurter BT, Aswad DW (2000) Co-operation between protein-acetylating and protein-methylating co-activators in transcriptional activation. Biochem Soc Trans 28(4), 415-418

Stubbs L (1992) Long-range walking techniques in positional cloning strategies. Mamm Genome 3, 127-142

Stubbs L, Rinchik EM, Goldberg E, Rudy B, Handel MA, Johnson D (1994) Clustering of six human 11p15 gene homologs within a 500-kb interval of proximal mouse chromosome 7. Genomics 24, 324-332

Sundarraj N, Pfeiffer SE (1973) Myelin basic protein arginine methyl transferase: wide distribution among both neurogenic and non-neurogenic tissues. Biochem Biophys Res Comm 52(3), 1039-1045

Sune C, Hayashi T, Liu Y, Lane SM, Young RA, Garcia-Blanco MA (1997) CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription. Mol Cell Biol 17(10), 6029-6039

Takahashi N, Roach A, Teplow DB, Prusiner SB, Hood L (1985) Cloning and characterization of the myelin basic protein gene from mouse: one gene can encode both 14 kd and 18.5 kd MBPs by alternate use of exons. Cell. 1985 42(1), 139-148

Talbot, WS, Schier AF (1999) Positional cloning of mutated zebrafish genes. Methods Cell Biol 60, 259-286

Tang J, Gary JD, Clarke S, Herschman HR (1998) PRMT3, a type I protein arginine *N*-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity, and regulation. J Biol Chem 273(27), 16935-16945

Tang J, Kao PN, Herschman HR (2000) Protein-arginine methyltransferase I, the predominant protein-arginine methyltransferase in cells, interacts with and is regulated by interleukin enhancer-binding factor 3. J Biol Chem 275(26), 19866-19876

Thomas JW, LaMantia C, Magnuson T (1998) X-ray-induced mutations in mouse embryonic stem cells. Proc Natl Acad Sci USA 95, 1114-1119

Threadgill DW, Dlugosz AA, Hansen LA, Tennenbaum T, Lichti U, Yee D, LaMantia C, Mourton T, Herrup K, Harris RC, et al (1995) Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. Science 269(5221), 230-4

Uberbacher EC, Mural RJ (1991) Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. Proc Natl Acad Sci USA 88(24), 11261-5

Valentini SR, Weiss VH, Silver PA (1999) Arginine methylation and binding of Hrp1p to the efficiency element for mRNA 3'-end formation. RNA 5, 272-280

Walkowicz M, Ji Y, Ren X, Horsthemke B, Russell L, Johnson D, Rinchik E, Nicholls R, Stubbs L (1999) Molecular characterization of radiation- and chemically induced mutations associated with neuromuscular tremors, runting, juvenile lethality, and sperm defects in *jdf2* mice. Mamm Genome 10, 870-878

Weiss VH, McBride AE, Soriano MA, Filman DJ, Silver PA, Hogle JM (2000) The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. Nat Struct Biol 7(12), 1165-1171

Whitaker JN (1998) Myelin basic protein in cerebrospinal fluid and other body fluids. Multiple Sclerosis 4, 16-21

Whitaker JN, Layton BA, Bartolucci AA, Mitchell GW, Bashir K, Goodwin J, Kachelhofer RD (1999) Urinary melin basic protein-like material in patients with multiple sclerosis during interferon beta-1b treatment. Arch Neurol 36, 687-691

Whitman S, Wang X, Shalaby R, Shtivelman E (2000) Alternatively spliced products CC3 and TC3 have opposing effects on apoptosis. Molecular and Cellular Biology 20(2), 583-593

Williams KR, Stone KL, LoPresti MB, Merill BM, Planck SR (1985) Amino acid sequence of the UP1 calf thymus helix-destabilizing protein and its homology to an analogous protein from mouse myeloma. Proc Natl Acad Sci USA 82(17), 5666-70

Wu M, Rinchik EM, Johnson DK (2000) An integrated deletion and physical map encompassing l71rl, a chromosome 7 locus required for peri-implantation survival in the mouse. Genomics 67, 228-231

Xiao H, Tao Y, Greenblatt J, Roeder RG (1998) A cofactor, TIP30, specifically enhances HIV-1, Tat-activated transcription. Proc Natl Acad Sci USA 95, 2146-2151

Xiao H, Palhan V, Yang Y, Roeder RG (2000) TIP30 has an intrinsic kinase activity required for up-regulation of a subset of apoptotic genes. EMBO J 19(5), 956-963

Yankulov K, Bentley D (1998) Transcriptional control: Tat cofactors and transcriptional elongation. Current Biol 8, R447-R449

You Y, Bergstrom R, Klemm M, Lederman B, Nelson H, Ticknor C, Jaenisch R, Schimenti J (1997) Chromosomal deletion complexes in mice by radiation of embryonic stem cells. Nat Genet 15(3), 285-288

Young PR, Waickus CM (1987) Time dependence of the methylation of myelin basic protein from bovine brain; evidence for protein-methylarginine demethylation. Biochem and Biophys Res Comm 142(1), 200-204

Young PR, Waickus CM (1988) Purification and kinetic mechanism of *S*-adenosylmethionine: myelin basic protein methyltransferase from bovine brain. Biochem J 250, 221-226

Young PR, Vacante DA, Waickus CM (1987) Mechanism of the interaction between myelin basic protein and the myelin membrane; the role of arginine methylation. Biochem Biophys Res Commun 145(3), 1112-1118

Zand R, Li MX, Jin X, Lubman D (1998) Determination of the sites of posttranslational modifications in the charge isomers of bovine myelin basic protein by capillary electrophoresis-mass spectroscopy. Biochemistry 37(8), 2441-2449

Zhang X, Zhou L, Cheng X (2000) Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. EMBO J 19(14), 3509-3519

Zhou Q, Sharp PA (1995) Novel mechanism and factor for regulation by HIV-1 Tat. EMBO J 14(2), 321-328

Zhou Q, Chen D, Pierstorff E, Luo K (1998) Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. EMBO J 17(13), 3681-3691

Zobel-Thropp P, Gary JD, Clarke S (1998) δ-*N*-methylarginine is a novel posttranslational modification of arginine residues in yeast proteins. J Biol Chem 273(45), 29283-29286

APPENDIX

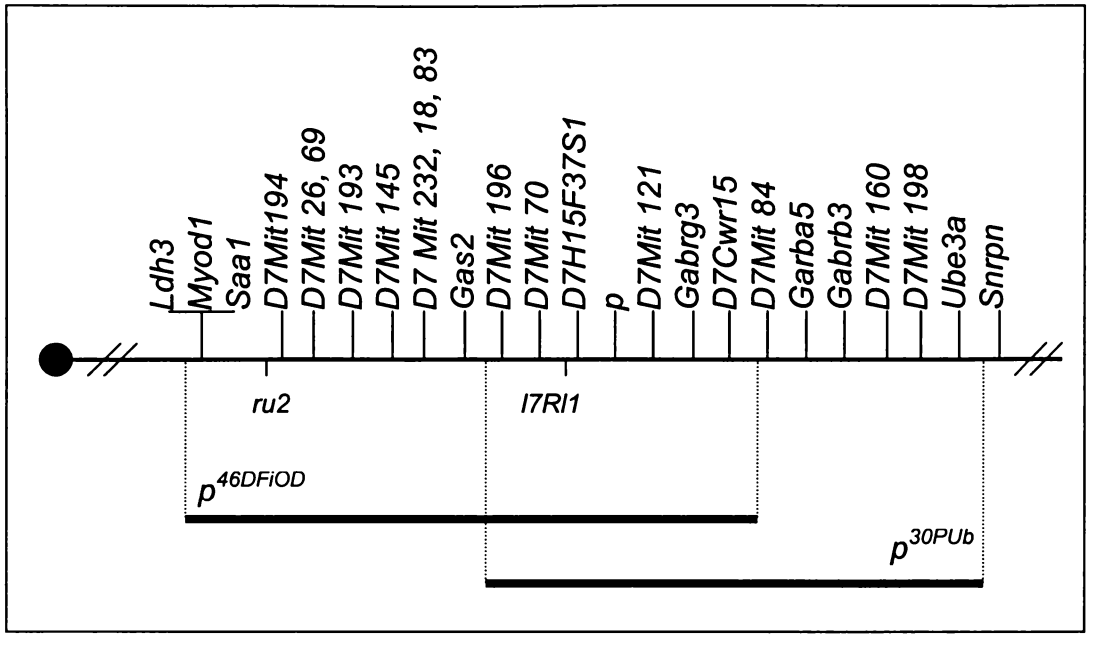


Figure II-1. The Limits of the *p*-Deletion Complex. Mouse chromosome 7 is represented by the central horizontal line, with the centromere to the left and the telomere to the right. Molecular markers are listed directly above the chromosome and phenotypes that have been mapped but not molecularly characterized are listed directly below the chromosome. *Ldh3*, *Myod1*, *Saa1*, and *ru2* are loci that cannot be ordered relative to the centromere. The heavy, shorter horizontal lines below the chromosome represent two of the *p* deletions. $p^{46DFiOD}$ is the most proximally extending deletion and p^{30PUb} is the most distally extending deletion. Their proximal and distal breakpoints, respectively, define the limits of the *p*-deletion complex. This figure is adapted from Dhar MS, Johnson DK (1997) A microsatellite map of the *pink-eyed dilution* (*p*) deletion complex in mouse Chromosome 7. Mamm Genome 8, 143-145.

Figure II-2. Mutagenesis Scheme Using the $p^{46DFiOD}$ Deletion to Select New Recessive Mutations. The mutagenized chromosome has two morphological markers, *ru2* and *p*. The mutagenized G_0 male animal is crossed to an animal that is homozygous wild type at both the *p* and *ru2* loci to produce phenotypically wild-type G_1 offspring. G_1 animals are crossed to animals carrying the $p^{46DFiOD}$ deletion opposite *ru2* (in this case). The G_2 offspring fall into three classes: wild-type (which represents two genotypes), the *ru2* carrier class, and the *ru2 p* test class. This figure is adapted from Rinchik EM, Carpenter DA, Handel MA (1995) Pleiotropy in microdeletion syndromes: Neurological and spermatogenic abnormalities in mice homozygous for the p^{6H} deletion are likely due to dysfunction of a single gene. Proc Natl Acad Sci USA 92, 6394-6398.

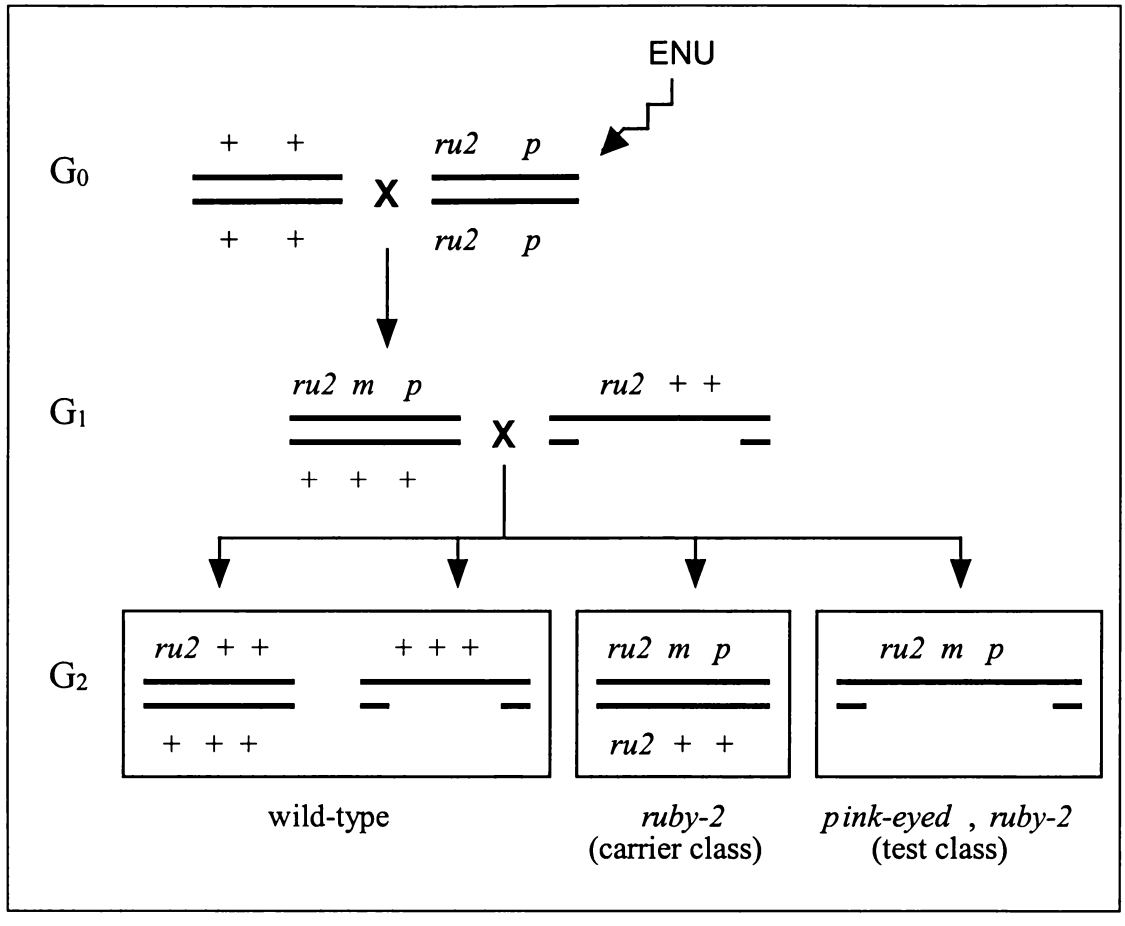


Figure II-2. Mutagenesis Scheme Using the $p^{46DFiOD}$ Deletion to Select New Recessive Mutations.

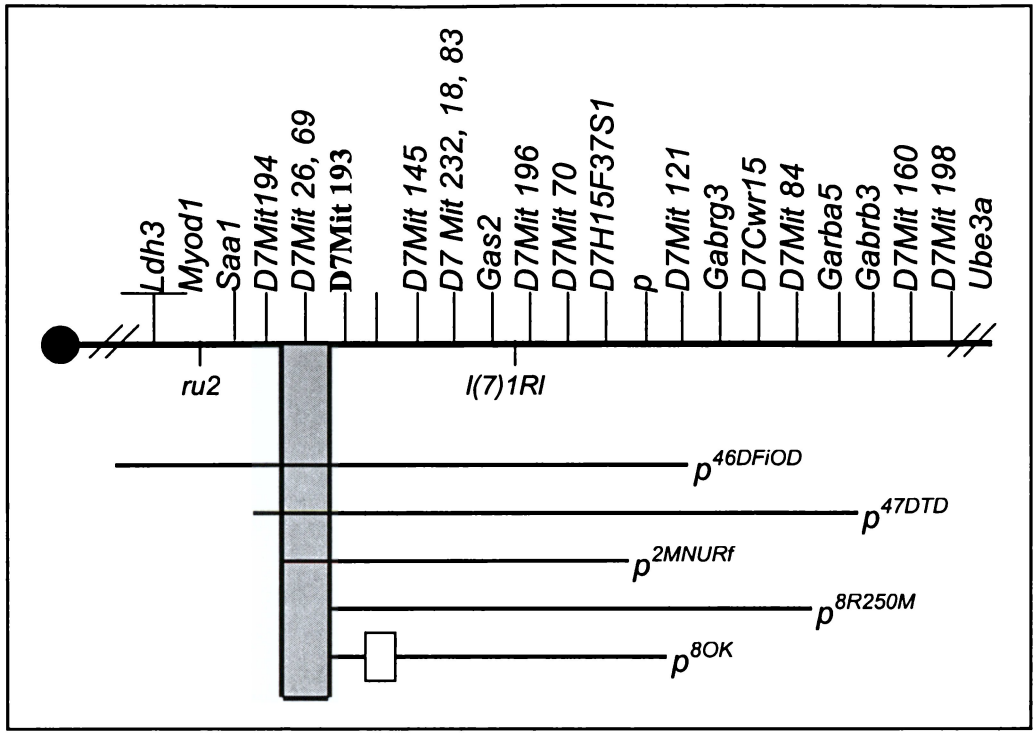


Figure III-1. Mapping the *psrt* Phenotype by *Trans*-Complementation Analysis. There are four *p* deletions that fail to complement the *psrt* phenotype: $p^{46DFiOD}$, p^{47DTD}, p^{2MNURf}, and $p^{3RD300H}$. All other deletions complement the phenotype. Therefore, the *psrt* minimal region, which is represented by the large gray rectangle, is defined as the deletion interval bordered proximally by the proximal breakpoint of p^{2MNURf} and bordered distally by the proximal breakpoints of p^{8R250M} and p^{8OK}, which cannot be resolved at this time. At the beginning of this project, the endpoints of the $p^{3RD300H}$ deletion had not been well characterized, so this deletion is not included in this map. *Ldh3*, *Myod1*, and *Saa1* are loci that cannot be ordered relative to the centromere. The white box on the p^{8OK} deletion indicates that it is a "skipper," or a noncontiguous deletion that is most likely a complex rearrangement.

Table IV-1. STSs Used in Physical Mapping.

Name of STS	Description of STS	Method	How Used
D7Mit26	microsatellite	PCR	BAC DNAs
D7Mit69	microsatellite	PCR	BAC DNAs
D7Mit70	microsatellite	PCR	BAC DNAs
D7Mit84	microsatellite	PCR	BAC DNAs
D7Mit145	microsatellite	PCR	BAC DNAs
D7Mit160	microsatellite	PCR	BAC DNAs
D7Mit193	microsatellite	PCR	BAC DNAs
D7Mit229	microsatellite	PCR	BAC DNAs
D7Mit230	microsatellite	PCR	BAC DNAs and deletion panel DNAs
RT 591/592	RT-PCR product (*Prmt3*)	PCR and Hybridization	BAC DNAs and BAC Southern blots
RT 595/596	RT-PCR product (*Prmt3*)	Hybridization	Deletion panel blots
B36n10/EcoRV/SP6	BAC end	Hybridization	BAC Southern blots
B36n10/EcoRV/T7	BAC end	Hybridization	BAC Southern blots
B45l4/Pvu/T7	BAC end	Hybridization	Deletion panel blots and BAC Southern blots
ORN 609/610	PCR product (same as B45l4/Pvu/T7)	PCR	BAC DNAs
B179d14/EcoRV/SP6	BAC end	Hybridization	Deletion panel blots and BAC Southern blots
B179d14/EcoRV/T7	BAC end	Hybridization	Deletion panel blots and BAC Southern blots
B36n10/B2A	Single Primer PCR product	Hybridization	Deletion panel blots
B179d14/B1B	Single Primer PCR product	Hybridization	Deletion panel blots

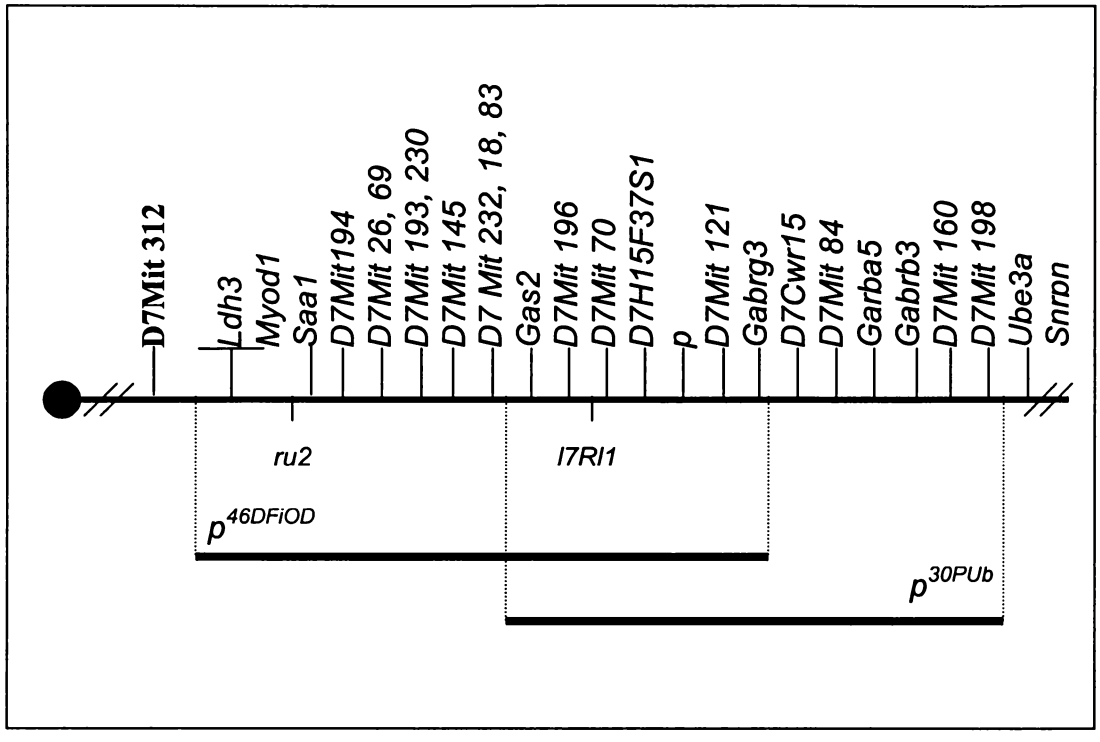


Figure IV-2. The *p*-Deletion Complex with New Microsatellite Markers. *D7Mit230* and *D7Mit312*, two markers that were mapped as part of this project, are included on this map. *D7Mit230* maps to the same deletion interval as *D7Mit193*, and *D7Mit312* maps outside the *p*-deletion complex. The two most proximally and distally extending of the *p*-deletions, $p^{46DFiOD}$ and p^{30PUb}, are shown. This figure is adapted from Dhar MS, Johnson DK (1997) A microsatellite map of the *pink-eyed dilution* (*p*) deletion complex in mouse Chromosome 7. Mamm Genome 8, 143-145.

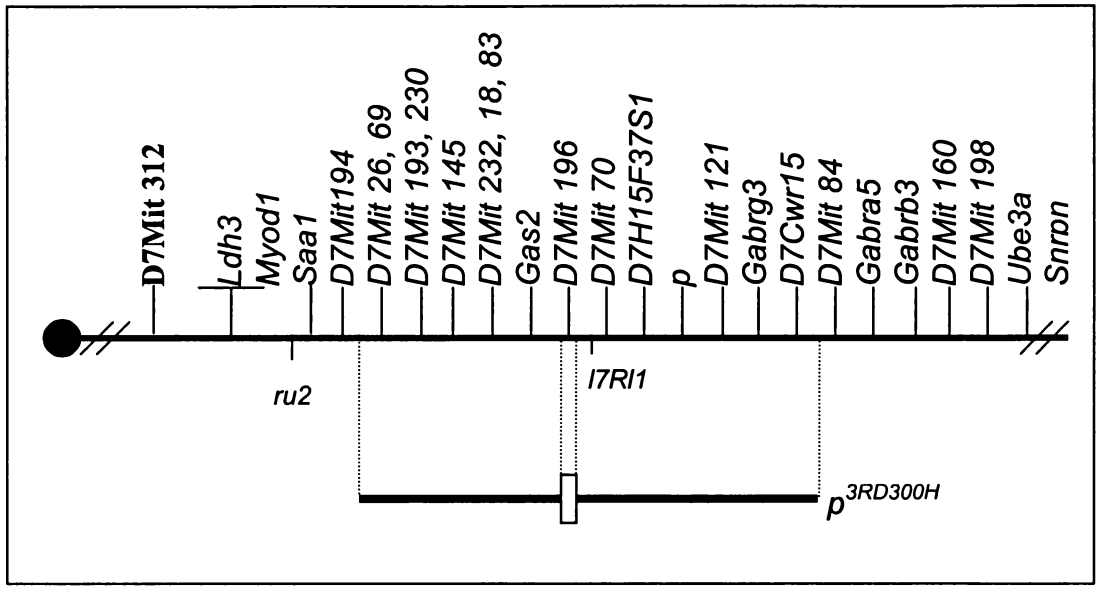


Figure IV-3. Defining the Breakpoints of the $p^{3RD300H}$ Deletion. PCR analysis indicates that three microsatellite markers map within the $p^{3RD300H}$ deletion: *D7Mit193*, *D7Mit230*, and *D7Mit145*. $p^{3RD300H}$ does not delete five other microsatellites that were tested: *D7Mit26*, *D7Mit69*, *D7Mit70*, *D7Mit84*, and *D7Mit160*. These data indicate that $p^{3RD300H}$ is a noncontiguous deletion (or a "skipper") since it does not delete *D7Mit70*. Previous studies indicate that the $p^{3RD300H}$ deletion encompasses three molecular markers: *p*, *Gas2*, and *D7H15F37S1* (Johnson *et al.*, 1995). This places the proximal breakpoint of $p^{3RD300H}$ between *D7Mit26* and *D7Mit69*, which cannot be resolved at this time, and the distal breakpoint between *D7Mit84* and *Gabra5*.

Table IV-4. Physical Mapping Data from BACs.

BAC	*D7 Mit 145*	*D7 Mit 193*	*D7 Mit 194*	*D7 Mit 229*	*D7 Mit 230*	RT 591/ 592	RT 595/ 596	ORN 607/608 (B179d14 end)	ORN 609/610 (B45l4/T7)	B36n10/ B2A	B36n10/ SP6	B36n10/ T7	B179d14 end	B179d14/ B1B	B207 p17/SP6	B211 i3/T7
B36n10		+								+	+	+				
B45l4		+							+			+				
B179d14		+				+	+	+	+	+	+	+	+	+		
B207p17	+														+	
B211i3				+	+											+
B212f2	+															
B47g18			+													
B47g22			+													
B142k10			+													
B142k22			+													
B178n12			+													
B242g12		+						+	+							
B279l14																
B288k21		+							+							
B315k22						+		+								
B329d6		+				+		+	+							
B333e23				+		+		+								
B344l6		+							+							
B354k11						+										
B372f7						+			+							
B382o13						+			+							
B434g10		+		+		+		+	+							
B462a15		+							+							

Note: Molecular markers in the table are not listed in their order along the chromosome. A (+) indicates the marker is present on the BAC.

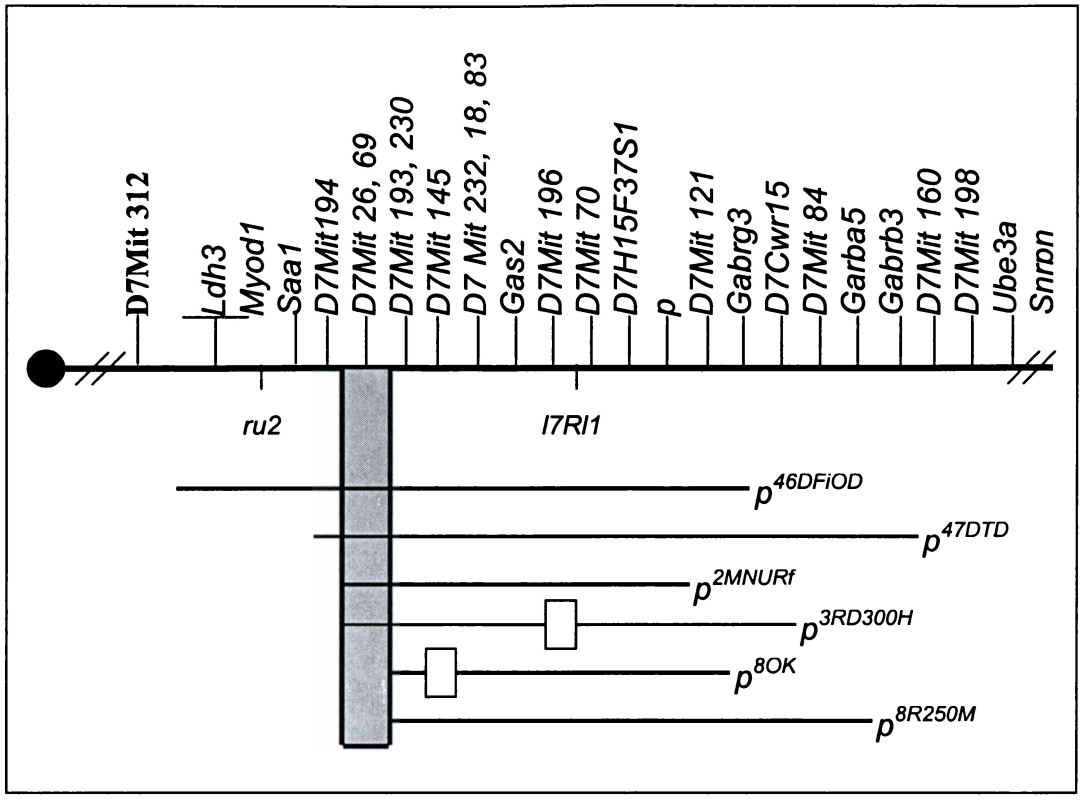


Figure IV-5. The Revised *psrt* Minimal Deletion Interval. The fine-mapping of the $p^{3RD300H}$ deletion enables the map of the minimal deletion interval to be revised to include the extents of all four *p* deletions that fail to complement the *psrt* phenotype: $p^{46DFiOD}$, p^{47DTD}, p^{2MNURf}, and $p^{3RD300H}$. All other deletions that have been tested complement the phenotype. Therefore, the *psrt* minimal region, which is represented by the large gray rectangle, is now defined as the deletion interval bordered proximally by the proximal breakpoints of both p^{2MNURf} and $p^{3RD300H}$, which cannot be resolved at this time. The region is bordered distally by the proximal breakpoints of p^{8OK} and p^{8R250M}, which also cannot be resolved at this time.

Table IV-6. Insert Sizes and *Not*I Restriction Data for Selected BACs.

BAC	Type of Library	Screened Using	Size of Insert	*Not*I Restriction Sites
B36n10	PCR pools	*D7Mit193*	120 kb	1
B45l4	PCR pools	*D7Mit193*	120 kb	1
B179d14	PCR pools	*D7Mit193*	150 kb	1
B207p17	PCR pools	*D7Mit145*	280 kb	2
B211i3	PCR pools	*D7Mit230*	125 kb	1
B212f2	PCR pools	*D7Mit145*	190 kb	0

Figure IV-7. Physical Map of the *psrt* Region of Mouse Chromosome 7. The centromere of the chromosome is to the left of the diagram, and the telomere is to the right. The relevant *p* deletions are drawn above the chromosome, with the line representing the area deleted. BACs are drawn below the chromosome and are labeled with the BAC name and size, if known. The *psrt* minimal interval is represented by the bold vertical lines and is labeled at the bottom of the figure. Molecular markers relevant to this study are represented as colored dots. Grey dots are microsatellite markers, with light grey representing those that have been previously mapped and dark grey representing those mapped as part of this work. Blue dots represent BAC ends, yellow dots are STSs generated using single primer PCR, and red dots are expressed sequences generated from RT-PCR. This map is not drawn to scale.

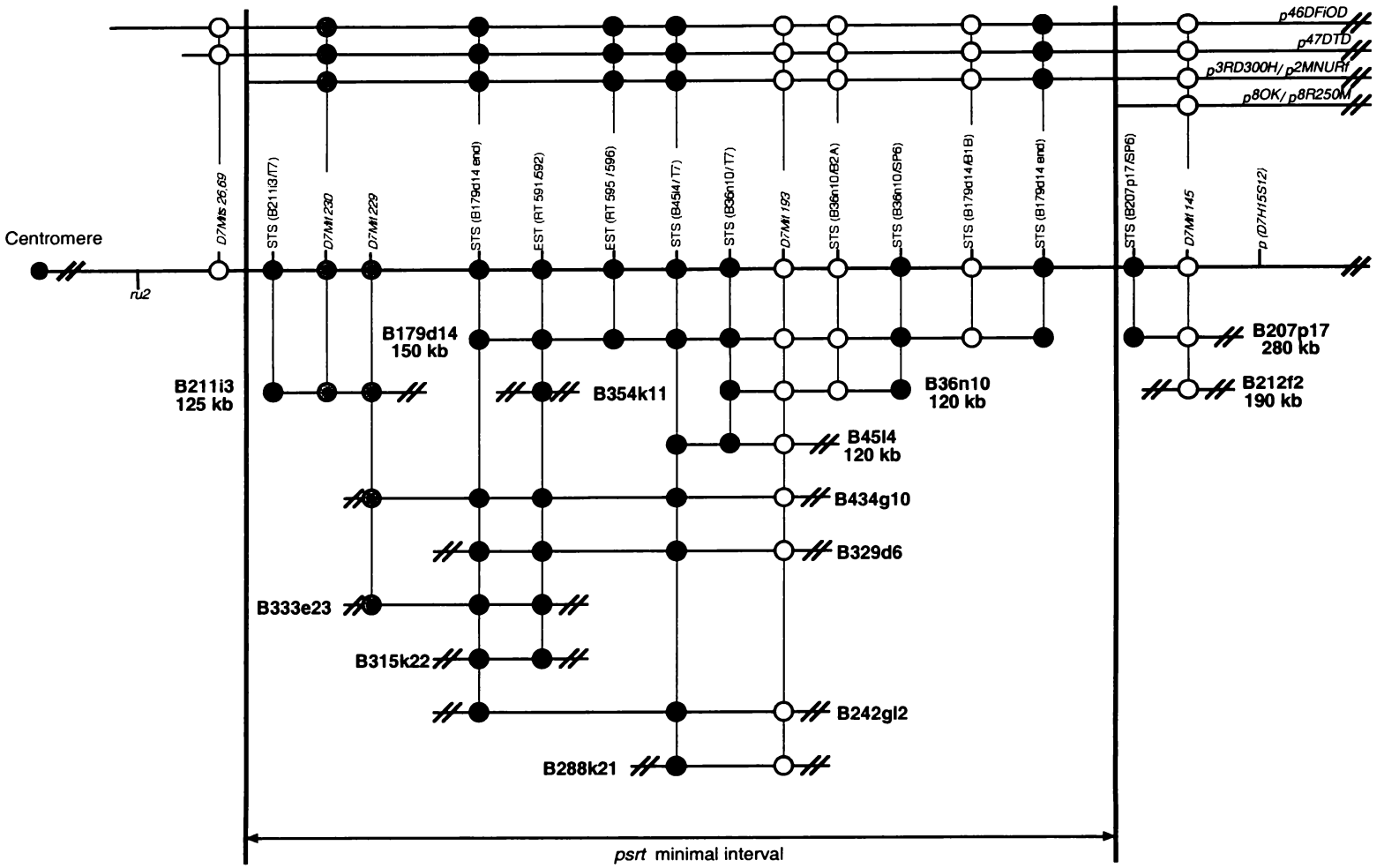


Figure IV-7: Physical Mapping in the *psrt* Region of Mouse Chromosome 7

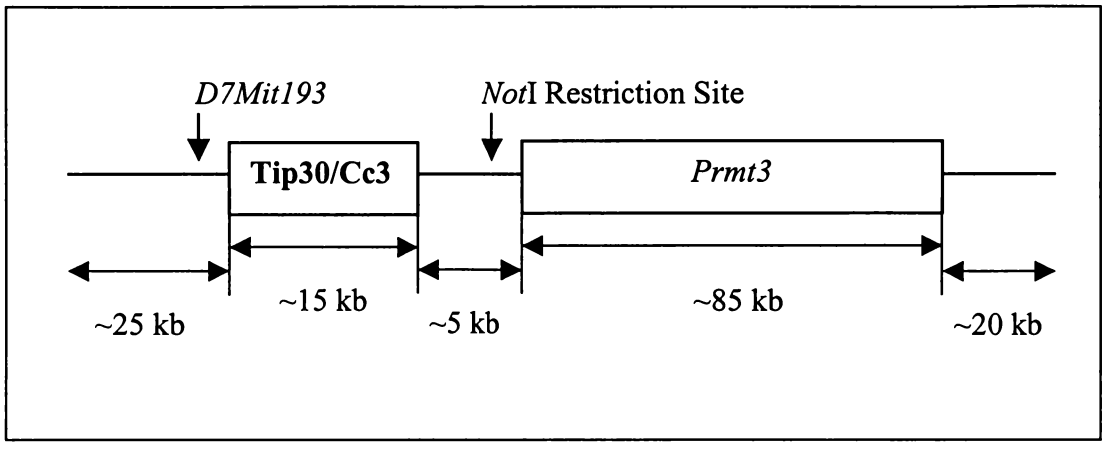


Figure V-1. Structure of the Insert of BAC B179d14. The *Tip30/Cc3* and *Prmt3* genes are transcribed in the same direction. The *Not*I restriction site is in the CpG island near the 5' end of *Prmt3*, and *D7Mit193* is near the 5' end of *Tip30/Cc3*. The length of the BAC insert is approximately 150 kb.

Table V-2. Exon and Intron Sizes for *Tip30/Cc3*.

EXON #	INTRON #	SIZE (bp)
1		283+
	1	2300
2		108
	2	8000
3		138
	3	1600
4		62
	4	660
5		226

Table V-3. Exon and Intron Sizes for *Prmt3*.

EXON #	INTRON #	SIZE (bp)
1		105+
	1	80
2		133
	2	1600
3		83
	3	730
4		50
	4	770
5		103
	5	4000
6		160
	6	1000
7		142
	7	4600
8		66
	8	6300
9		122
	9	8100
10		10
	10	11400
11		79
	11	8500
12		188
	12	1800
13		87
	13	19100
14		51
	14	200
15		88
	15	8400
16		109
	16	776

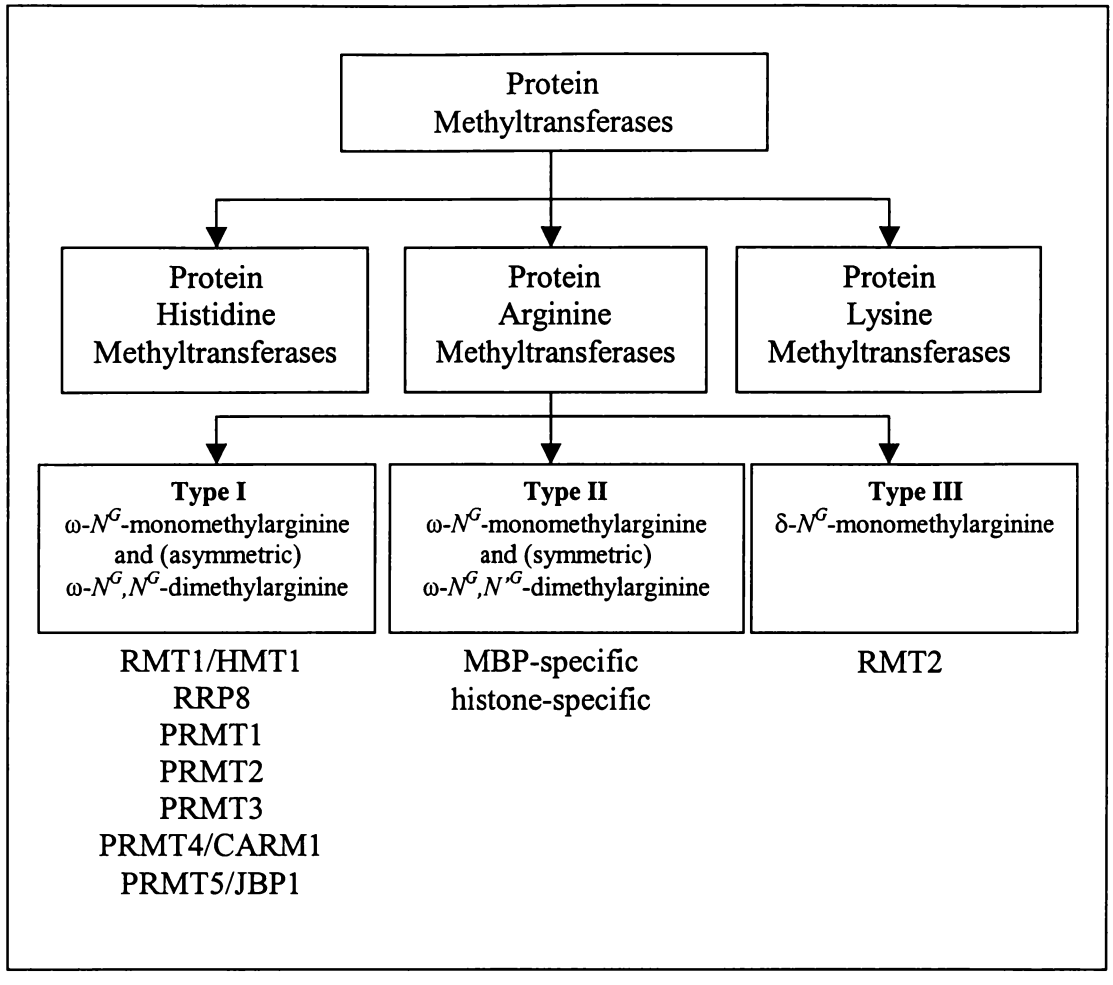


Figure VI-I. Protein Methyltransferases.

Table VII-1. Primers Used for Standard PCR, RT-PCR and RACE.

Primer Name	Primer Sequence (5' → 3')	Gene	Corresponding bp of cDNA
ORN 644	gaa gtt gaa gca cat gga gg	*Prmt3*	554-573
ORN 645	atc ttc tgc aga tcc tca cg	*Prmt3*	712-693
ORN 646	gtc gat gtt gga ttc tgt cc	*Prmt3*	1166-1185
ORN 647	aat gtc agg gta gac cga tc	*Prmt3*	1238-1219
ORN 648	cca tcc cgg tca ctg cc	*Prmt3*	1751-1735
ORN 663	tgc cca aca gtt aac aac cc	*Tip30/Cc3*	1-20
ORN 664	gat ccc tgt ttg caa cct ac	*Tip30/Cc3*	1059-1039
ORN 704	cgc gct tgg tcg g	*Prmt3*	(-18) - (-6)
ORN 723	gag aag ctg gat ggc tat g	*Tip30/Cc3*	305-323
ORN 726	gat ttg gtg aca ggg cac tg	*Tip30/Cc3*	323-305
ORN 727	cta ctg tgt gac agg caa g	*Tip30/Cc3*	596-614
ORN 728	ctt gcc tgt cac aca gta g	*Tip30/Cc3*	614-596
ORN 772	gcg cgt gag gat ctg cag aag atg	*Prmt3*	632-609

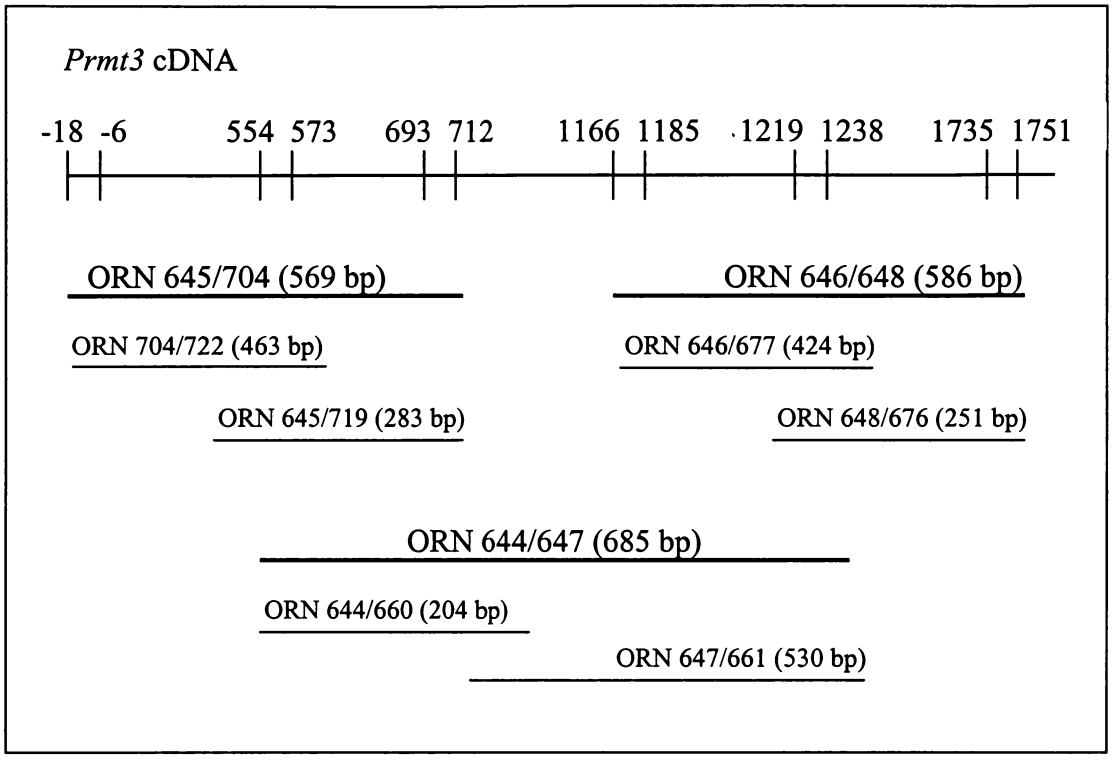
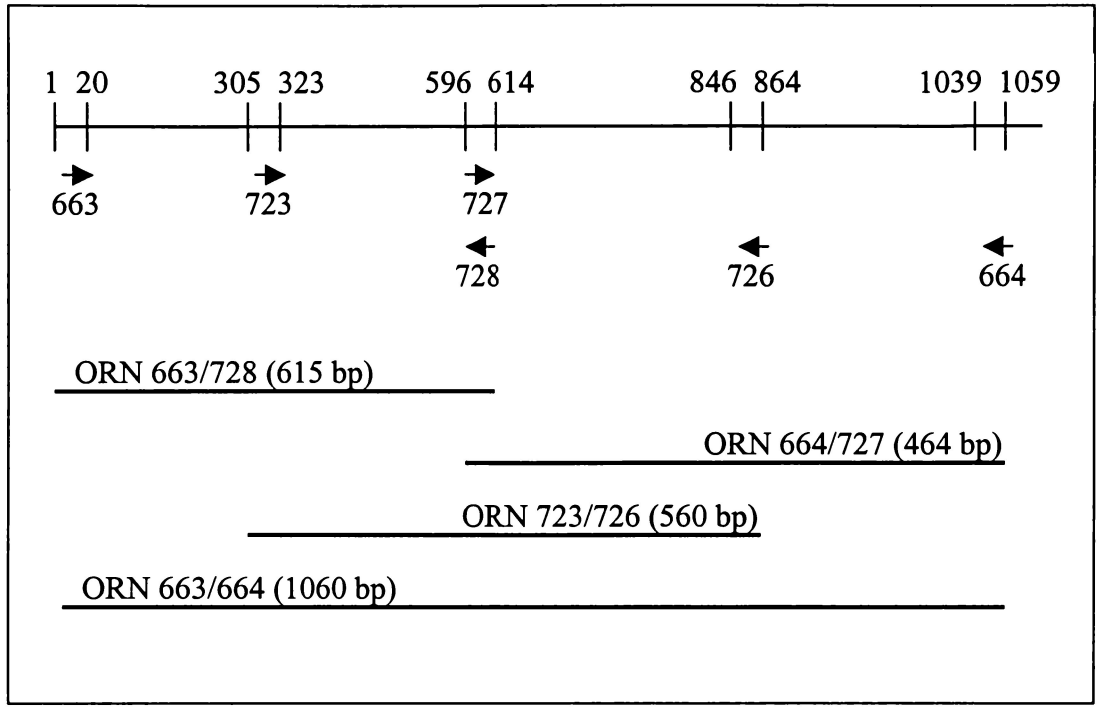


Figure VII-2. *Prmt3* cDNA Fragments Used for Heteroduplex Analysis. The *Prmt3* cDNA is represented by the long horizontal line at the top of the figure, and the numbers refer to the nucleotides corresponding to some of the PCR primers. (The start codon is at position 1.) The shorter, horizontal lines in lighter print are the smaller RT-PCR amplification fragments analyzed using TGCE. These fragments range in size from 204 bp to 530 bp. The darker horizontal lines below the cDNA are the original RT-PCR fragments used for sequence analysis, which range in size from 569 to 685 bp.

Figure VII-3. RT-PCR Analysis of the *Tip30/Cc3* cDNA. (a) The cDNA is represented by the numbered horizontal line near the top of the figure, numbered so that bp 1 is the first nucleotide of the start codon. Three sets of primers, represented by the numbered arrows, were used to amplify overlapping fragments for sequence analysis. The horizontal lines at the bottom of the figure represent the amplification fragments. (b) RT-PCR amplification products run on ethidium bromide-stained agarose gels. A single primer pair (ORN 663/664) was used to amplify a 1060 bp fragment representing the entire *Tip30/Cc3* coding sequence for expression analysis. The mutants and controls show no size difference. A 520 bp amplification fragment from *β-actin*, amplified using primers ORN 559/560, was the control.

(a)



(b)

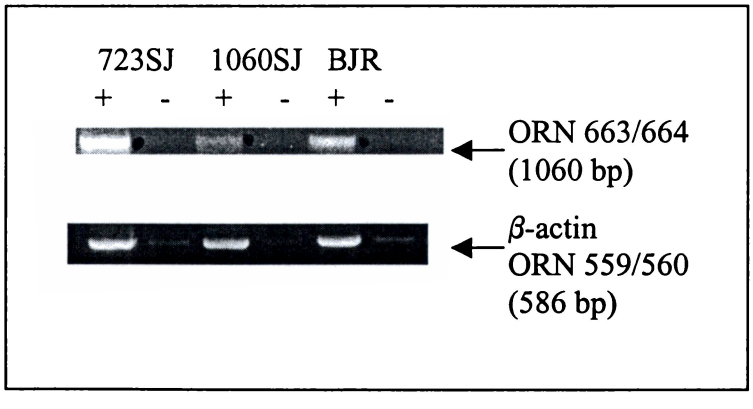


Figure VII-3. RT-PCR Analysis of the *Tip30/Cc3* cDNA.

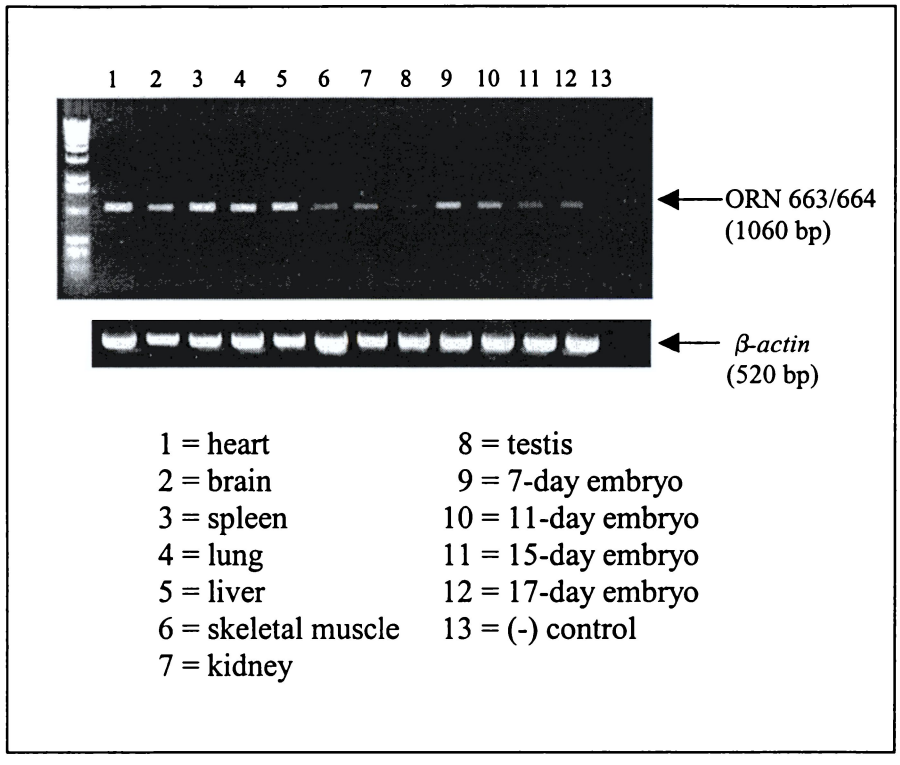
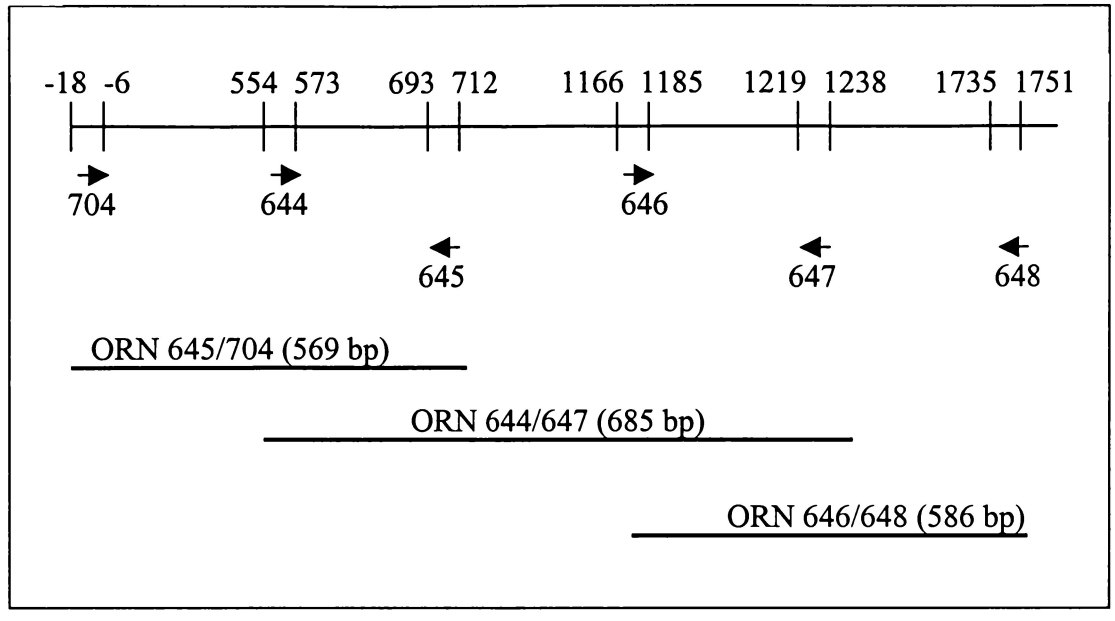


Figure VII-4. *Tip30/Cc3* Multiple Tissue Expression Analysis. Primers ORN 663/664 were used with cDNAs from the mouse multiple tissue panel to amplify a 1060 bp fragment of the *Tip30/Cc3* cDNA containing the entire coding sequence. A single band of the expected length was detected in all tissues analyzed. A 520 bp fragment from *β-actin*, amplified using primers ORN 559/560, was the control.

Figure VII-5. RT-PCR Analysis of the *Prmt3* cDNA. (a) The cDNA is represented by the numbered horizontal line near the top of the figure, numbered so that bp 1 is the first nucleotide of the start codon. Three sets of primers, represented by the numbered arrows, were used to amplify overlapping fragments for sequence analysis. The horizontal lines at the bottom of the figure represent the amplification fragments. (b) RT-PCR amplification products run on ethidium bromide-stained agarose gels. The mutants and controls show no size difference in the bands; however, the 723SJ cDNA produces no amplification product with the ORN 645/704 primers. A 520 bp amplification fragment from *β-actin*, amplified using primers ORN 559/560, was the control.

(a)



(b)

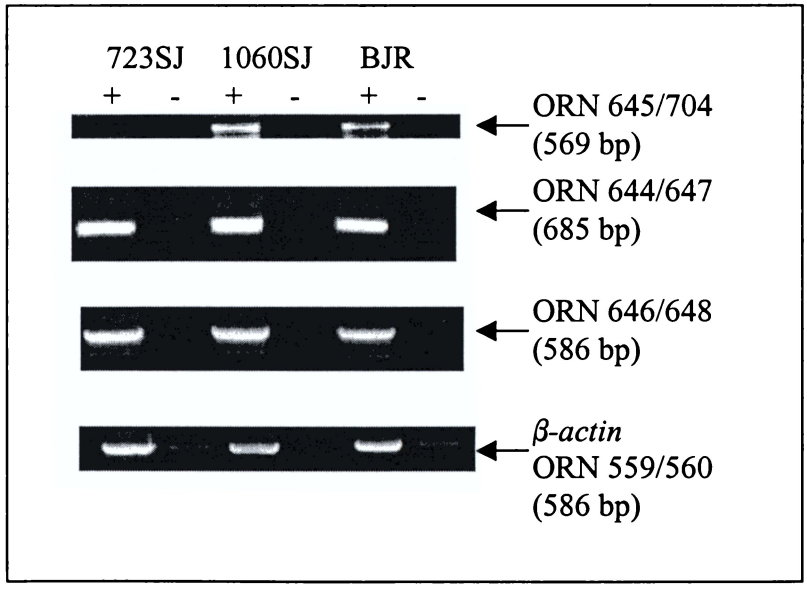


Figure VII-5. RT-PCR Analysis of the *Prmt3* cDNA.

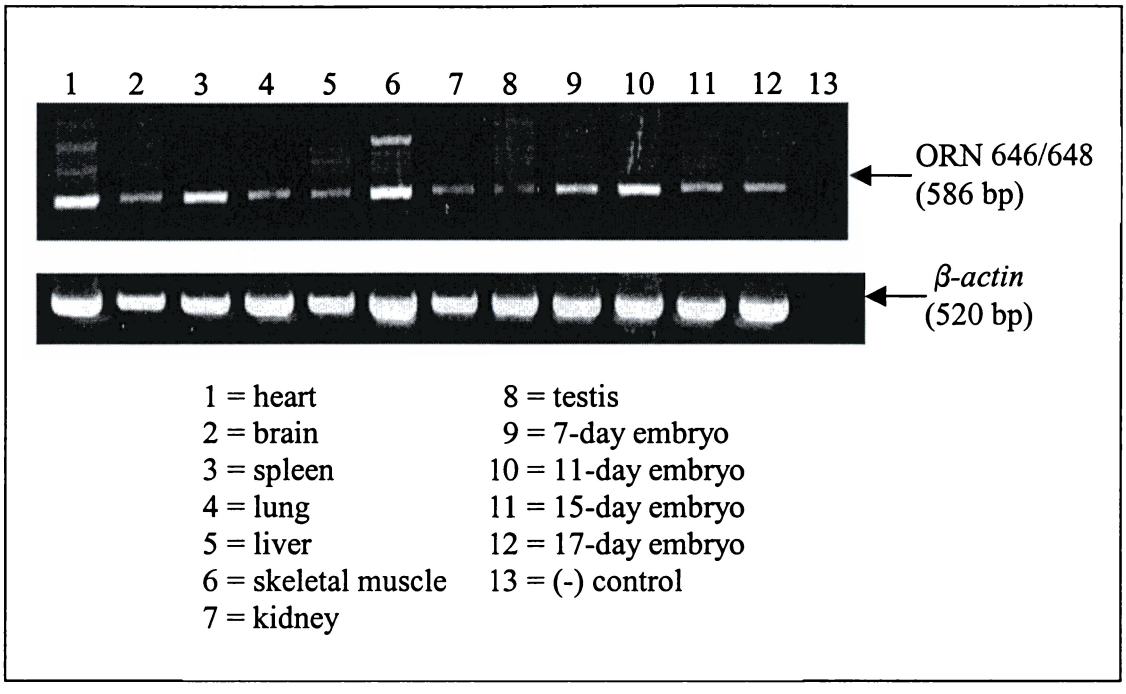


Figure VII-6. *Prmt3* Multiple Tissue Expression Analysis. Primers ORN 646/648 were used with cDNAs from the mouse multiple tissue panel to amplify a 586 bp fragment from the 3' end of the *Prmt3* cDNA. A band of the expected length was seen is all tissues analyzed, and a second, larger band is seen in skeletal muscle and possibly in heart. A 520 bp fragment from *β-actin*, amplified using primers ORN 559/560, was the control.

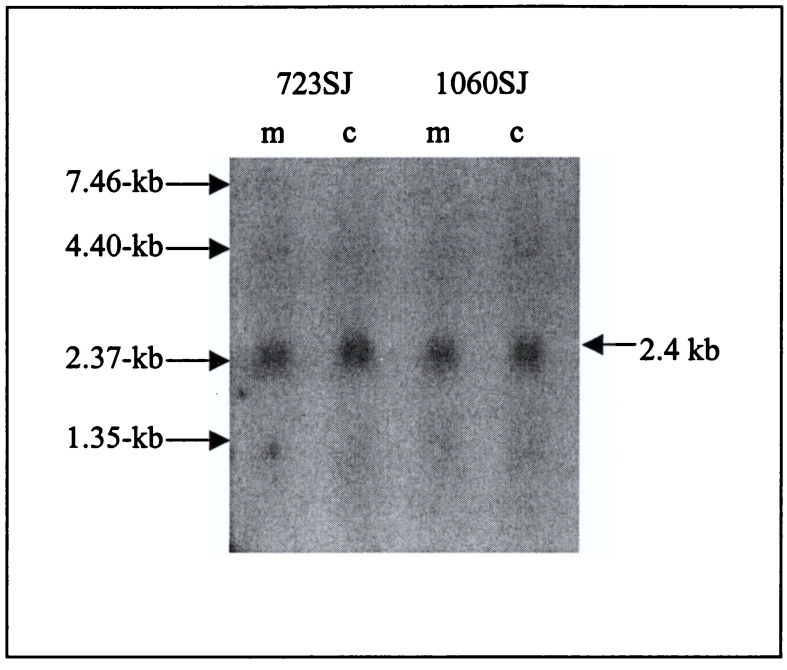


Figure VII-7. Northern Blot of Mutant and Control Brains Probed with a *Prmt3* Probe. This blot contains total RNAs from the brains of *psrt* mutants (m) and non-mutant littermate controls (c) from both the 723SJ (genotype *ru2* + + / *Df* [*ru2-p*]) and 1060SJ (genotype + + *p* / *Df* [*ru2-p*]) mutant stocks. *Df* [*ru2-p*] represents the $p^{46DFiOD}$ deletion. The blot was probed with a RT-PCR fragment representing the 5' end of the *Prmt3* cDNA (ORN 645/704). The 723SJ mutant transcript appears to be slightly shorter than the other transcripts.

Table VII-8. Restriction Data for DNA Used in the Genomic Southern Blot (Figure VII-9).

Enzyme	# restriction sites	Smallest fragment	Largest fragment
*BstX*I	13	39	4966
*Eco*RI	5	3287	5979
*Pst*I	20	29	3735
*Taq*I	9	992	5258

These data represent 25-kb of genomic DNA, from 5-kb upstream of the start codon of *Prmt3* to 20-kb downstream.

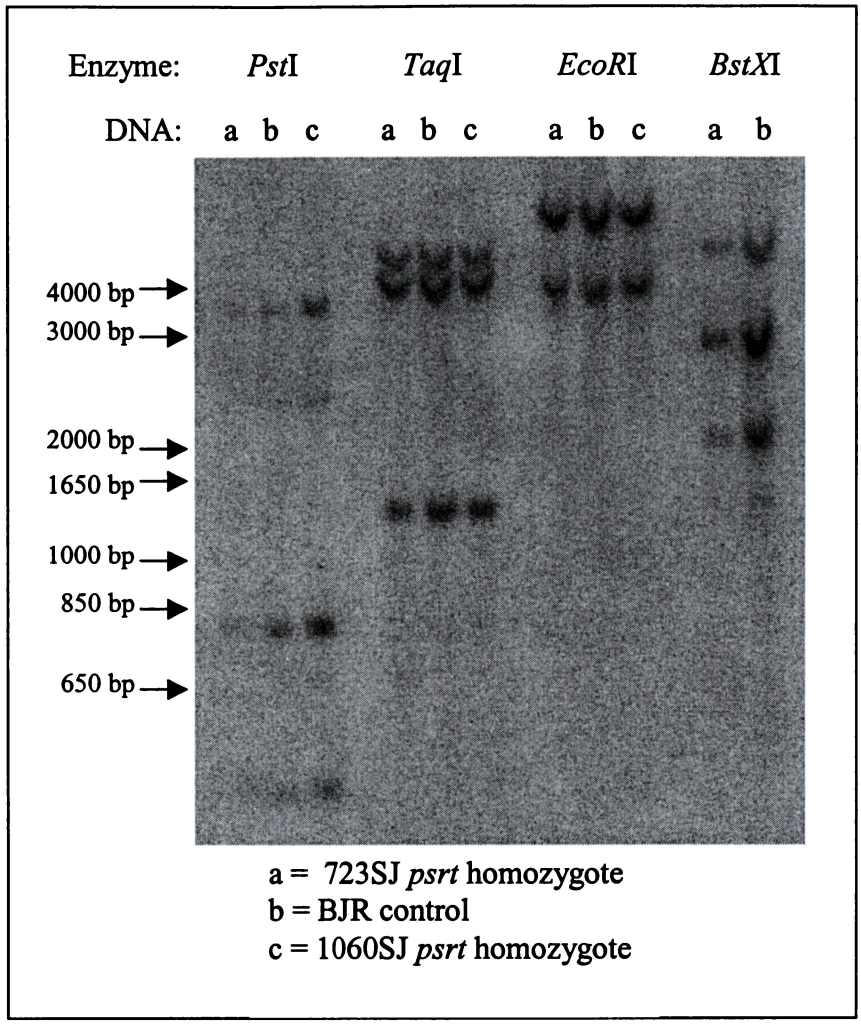


Figure VII-9. Autoradiograph from a Genomic Southern Blot of *psrt* mutants. The blot was probed with 5' end of the *prmt3* cDNA. Listed above the photograph are the enzymes used for digestion and a representation for the genomic DNA present in each lane. To the left of the photograph are arrows corresponding to the size markers from the 1-kb plus DNA ladder (not shown). The legend for the DNAs is below the photograph. This blot shows no evidence of a genomic rearrangement or deletion.

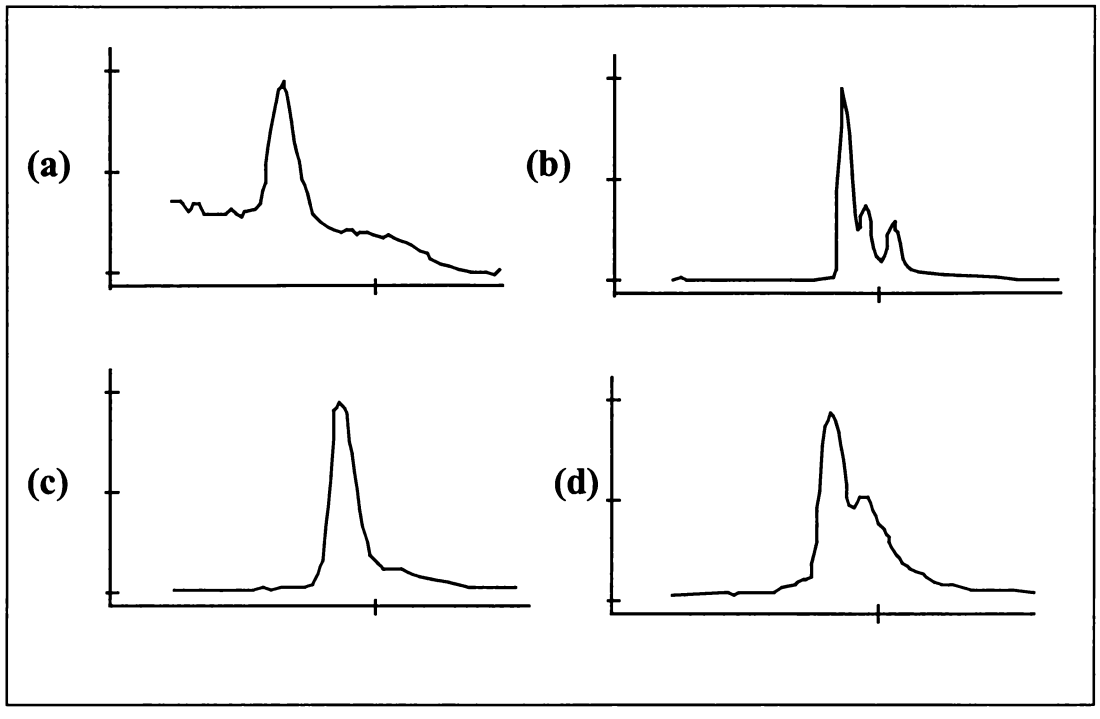


Figure VII-10. Electropherograms from the TGCE Analysis of the *psrt* Mutants. (a) The RT 646/677 fragment from the BJR control and the (b) 723SJ *psrt* homozygote mixed with the BJR control. There is a single large peak in electropherogram (a) and two smaller peaks to the right of the main peak in (b). (c) The RT 648/676 fragment from the BJR control and the (d) 1060SJ *psrt* homozygote. There is a single large peak in electropherogram (c) and a smaller peak that appears as a distinct "shoulder" to the right of the main peak in (b). The multiple peaks present in both (b) and (d) indicate a mismatch between the control and mutant cDNAs.

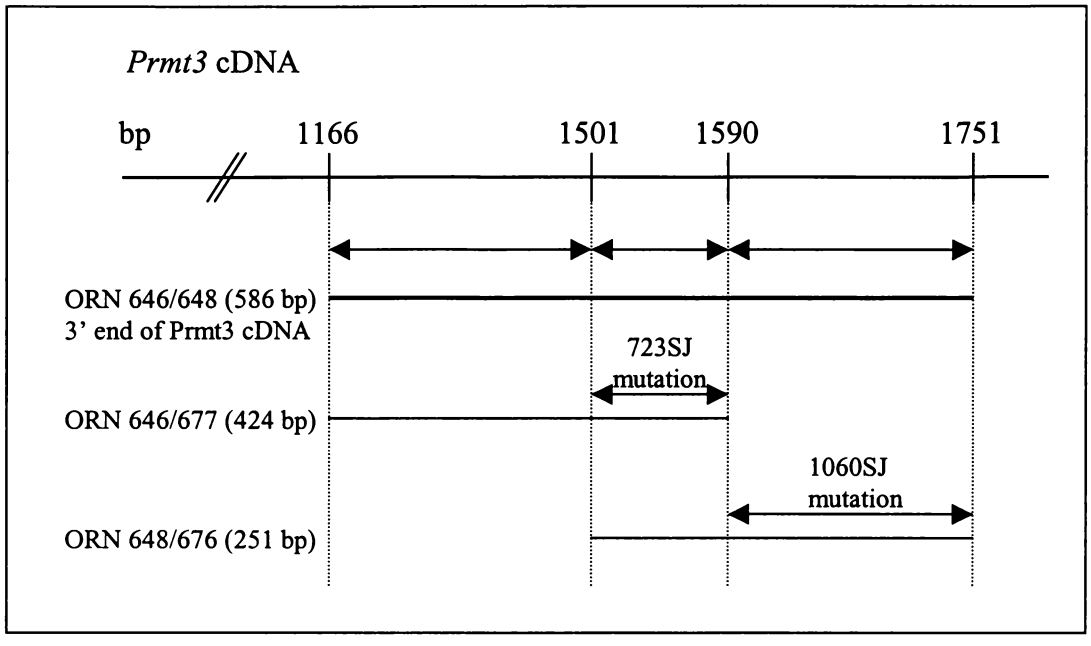


Figure VII-11. *Prmt3* RT-PCR Fragments Containing Mismatches. Mismatches were detected in both the ORN 646/677 and the ORN 648/676 fragments for the 723SJ strain, indicating that the mutation is most likely in the 89 bp of sequence that overlaps in those two fragments. A mismatch was detected only in the ORN 648/676 fragment in the 1060SJ strain, indicating that the mutation is most likely in the 162 bp of sequence that is contained only in the ORN 648/676 fragment.

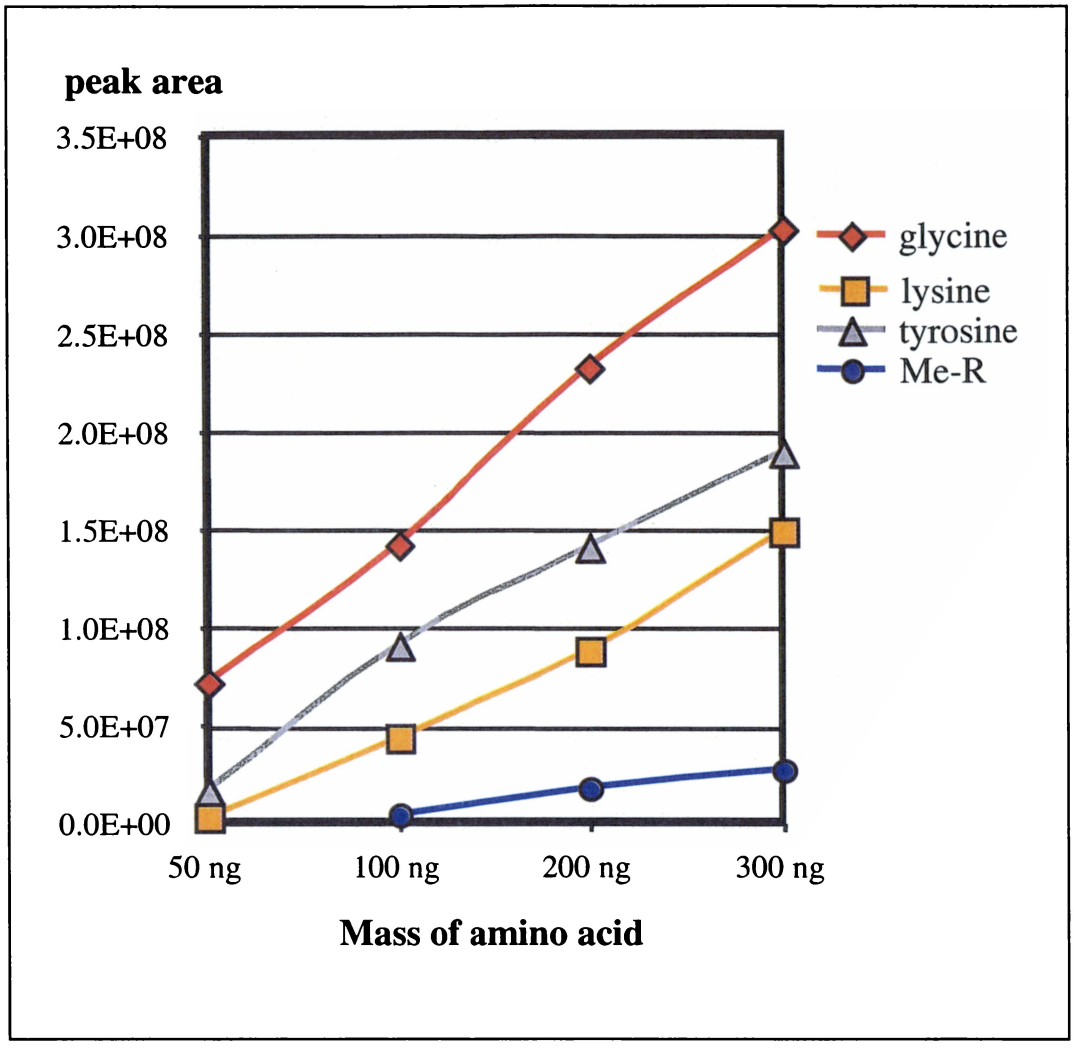


Figure VIII-1. Calibration Curves for Amino Acid Standards. The concentration of these four amino acids in a sample can be determined by extrapolation if the corresponding peak area is known.

Table VIII-2. Extrapolated Quantitation Data.

Sample	Lysine Peak Area	Lysine Quantity	Me-R Peak Area	Me-R Quantity	Tyrosine Peak Area	Tyrosine Quantity
psrt mutants[a]	41,414,848	~100 ng	N/A	N/A	13,242,394	~25 ng
non-*psrt* controls[b]	87,584,333	~200 ng	3,523,593	~100 ng	57,284,175	~75 ng

[a] The pooled mutant sample class included five homozygotes (genotype *ru2 m p / ru2 m p*) and two hemizygotes (genotype *ru2* m *p* / *Df*[*ru2-p*])

[b] The pooled control sample class included five wild-type animals (genotype *ru2* + + / + + *p*) and two non-carrier *ru2* animals (genotype *ru2* + + / *ru2* + +)

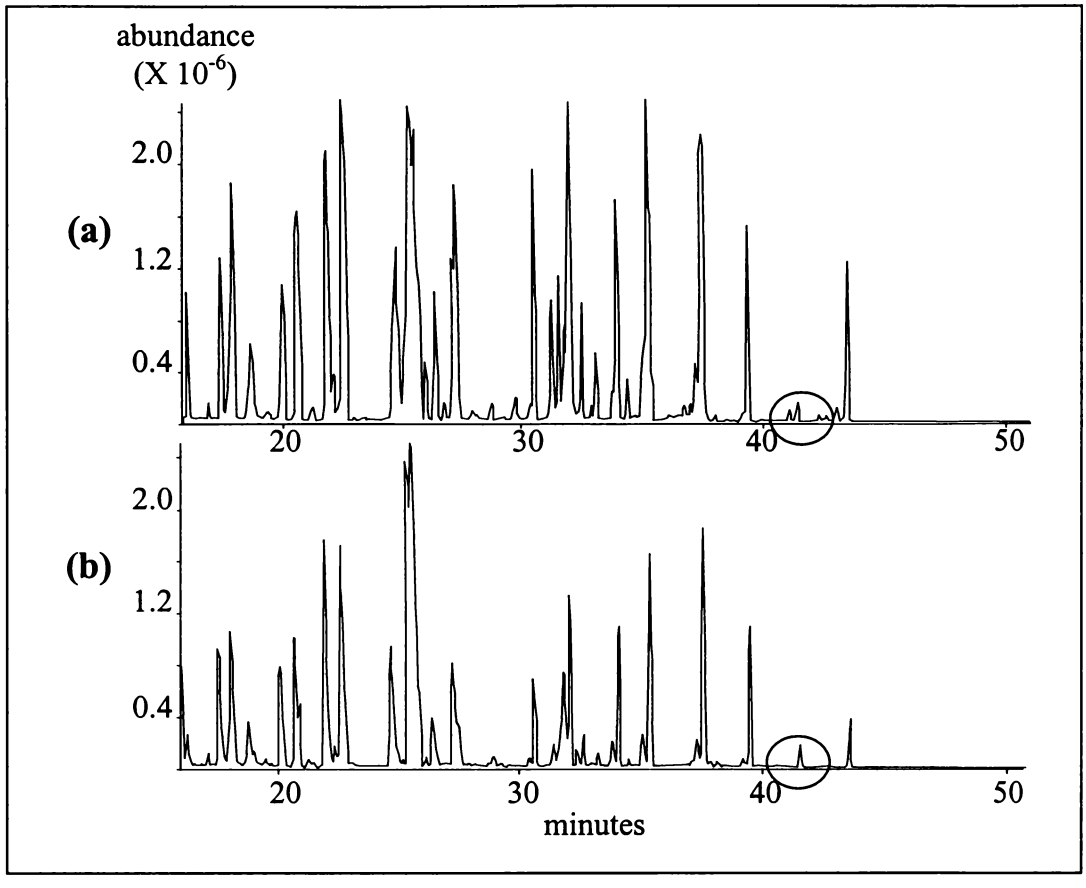


Figure VIII-3. Total Ion Chromatographs of MBP Hydrolysate from Control and Mutant Animals. Chromatographs from age-matched (a) 1060SJ control animals, and (b) 1060SJ *psrt* mutant animals. The genotypes of these animals are noted in Table VIII-2. The circled regions represent the area surrounding the Me-R peak, which is identified by its retention time and its mass spectrum. This peak is missing in the chromatograph from the mutant animal. An expanded view of the circled regions can be found in Figure VIII-4.

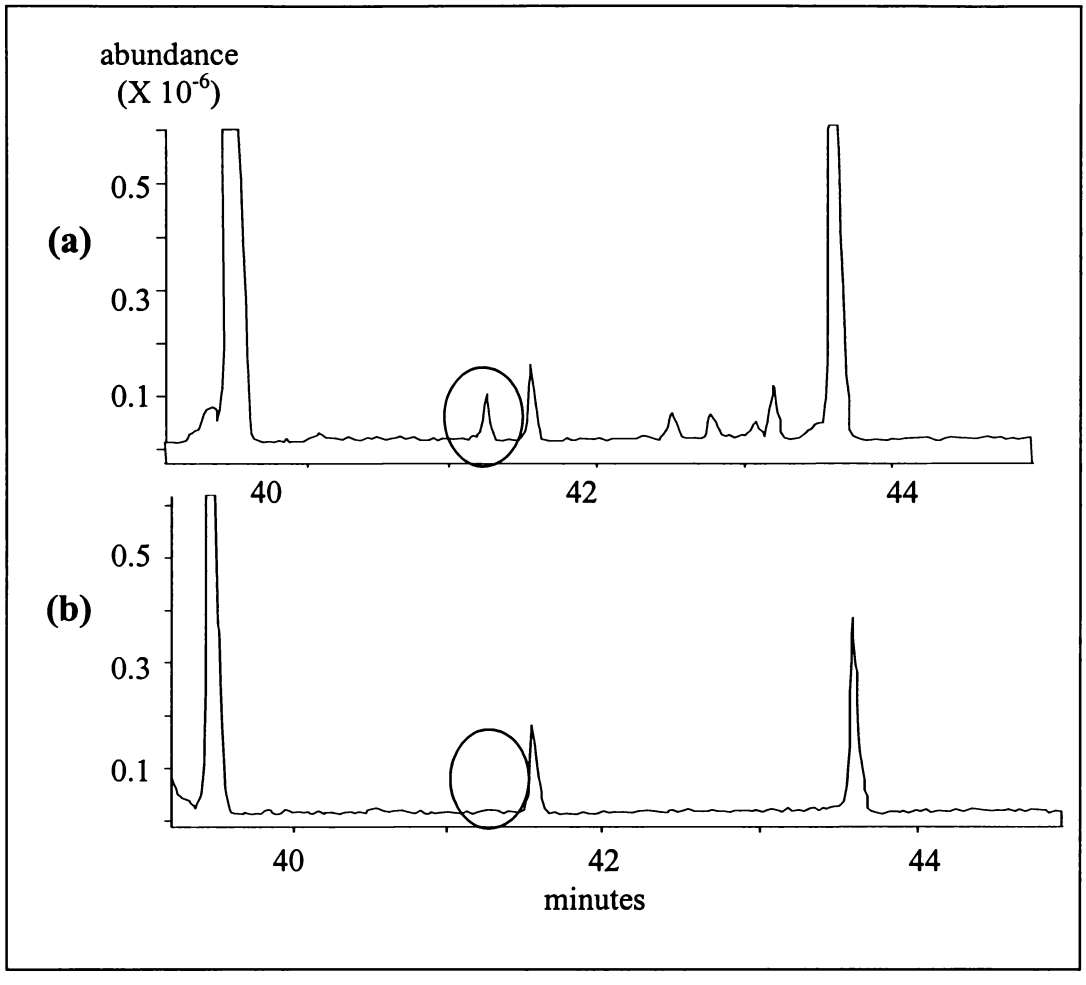


Figure VIII-4. Expanded View of the Me-R Region of the Total Ion Chromatographs of MBP Hydrolysate from Control and Mutant Animals. Chromatographs from age-matched (a) 1060SJ control animals, and (b) 1060SJ *psrt* mutant animals. The genotypes of these animals are noted in Table VIII-2. The circled regions represent the Me-R peak, which was identified by its retention time and its mass spectrum. This peak is not present in the chromatograph from the mutant animal.

VITA

Lisa Smith Webb was born in Chattanooga, Tennessee, on June 22, 1961, and was raised in Knoxville, Tennessee, where she attended schools in the Knox County Public School System. She graduated from Doyle High School in June 1979, and entered Maryville College in the fall of 1979, where she received the Bachelor of Arts in Chemistry in May 1985. She was awarded the Distinguished Achievement in Chemistry Award from Maryville College and the Outstanding Student Award from the American Institute of Chemists in the spring of 1985. She taught high school chemistry, physics, and physical science in the Gwinnett County (Georgia) Public School System for seven years and earned a Master of Education with a major in Secondary Science Education at the University of Georgia in August 1995. She entered the UT-ORNL Graduate School of Biomedical Sciences in August 1995 to pursue the Doctor of Philosophy in Biomedical Sciences with a concentration in Genetics, which was officially received in August 2001.

She received a National Institutes of Health Postdoctoral Fellowship and is currently working under the direction of Dr. Luanne L. Peters at The Jackson Laboratory in Bar Harbor, Maine.